Definition and estimation of intervention effects in complex systems: Gender equity in academia

by

Francis M. Abreu

A dissertation submitted to The Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy.

Baltimore, Maryland

April, 2016

Abstract

Even though gender equity in academia has been extensively studied, female faculty are still consistently hired at lower ranks, paid lower salaries and promoted less frequently than men.[1]

Previous work has focused on the individual faculty member as a study unit and, in most applications, on a single academic reward or representation outcome. However, existing approaches are insufficient to assess equity at institutional level for single-institution studies, from a causal inference perspective.

How do differential gender practices in awarding salaries and ranks affect institutional measures of prestige and investment? In this dissertation we developed a simulation-based approach to estimate and conduct inference for gender equity outcomes defined at institution level and investigate how gender disparities along individual careers contribute to institutional measures.

The statistical challenge in addressing these issues corresponds to the estimation of higher-level causal effects in complex systems for which only one observation of the outcomes of interest is available. The methods proposed combine faculty-level

models that describe the academic career with a university-level, i.e., an aggregate-level, definition of causal effect.

We applied these methods to simulated data based on an existing university. We found that the simulated institution does not deviate significantly from gender neutrality in terms of departures from the institution and total time in higher ranks for female faculty in 2005-2013. However, under a counterfactual gender-neutral scenario, the total compensation paid to female faculty over these 9 years would have been 2.8% higher (95% CI [1.2%, 4.4%]). The main determinant of this disparity is the significantly lower initial salaries for female faculty, with women earning 6.0% less on average at-hire than otherwise similar men.

This analysis aims to complement individual-level gender equity studies with an institutional perspective, to aid in the achievement of a more gender-neutral structure in academia. Furthermore, the methods proposed have wide applications to other complex systems and designs, such as health agencies networks, pharmaceutical market dynamics and transportation systems.

Advisors: Scott Zeger, Elizabeth Colantuoni

Committee: Joanne Katz, Janice Clements

Alternates: Richard Thompson, Janet Holbrook

Acknowledgments

To my advisors Scott Zeger and Elizabeth Colantuoni, without whom not a single page of this dissertation would have been written. Thank you for your infinite patience, for your invaluable advice, and for giving me excellent examples of what a researcher, a teacher, and a wonderful human being should be like. I want to be like you when I grow up.

To Tom Louis and Elizabeth Sugar, who got me started on my research path at Hopkins. Thank you for your knowledge, guidance, and patience while I miscounted eyes over and over again. Thank you for teaching me thoroughness and critical thinking and how to be a better writer and communicator.

To absolutely everyone at the Biostatistics department. There isn't a single day in which I don't think I made the best decision of my life coming here. You have been my family away from home.

To my thesis committee, Joanne Katz, Janice Clements, Richard Thompson and Janet Holbrook, thank you for taking the time to review my work. Your comments have made me a better researcher.

To my parents, without whom I wouldn't have made it here or through this PhD. You are the best parents anyone can ask for. This is why I chose you.

To my family, thank you for keeping me centered. Thank you for pushing me forward, and for reminding me that as long as I have you, I always have a place to come back to.

To friendships that span decades and continents, thank you for being with me at all times, just a click away.

Antonio and Cleiver, thank you for being a voice of reason and a safe place.

Yenny the Wise and Alyssa the Awesome, thank you for being with me from day one.

Pretty David and my baby Koala, thank you for being my light at dark times and my partners in crime.

José Miguel, there are really no words for all you did for me. Thank you for turning my life into a beautiful song. Cada vez que haga un brindis, brindaré por sólo una sobra de amor, por el bien de ambos y por mi Caracas mejor.

Dedication

To the strongest women I have ever known:

my aunts, Amalia and Enriqueta.

A las mujeres más fuertes que he conocido:

mis tías, Amalia y Enriqueta.

You are never given a wish without also being given the power to make it true. You may have to work for it, however.

Richard Bach

Contents

Abstract **ii**

Acknowledgments **iv**

List of Tables **xii**

List of Figures **xiii**

1 Introduction **1**

1.1 Gender equity in academia 5

1.1.1 Representation 8

1.1.2 Academic rewards 9

1.2 Data and confidentiality 12

1.3 Gender equity at University X: descriptive analyses 15

2 Literature review: statistical approaches to asses gender equity **19**

2.1 Existing methods 21

2.1.1 Matching . . . 21
2.1.2 Compa-ratios . . . 22
2.1.3 Multivariable regression . . . 22
2.1.4 Decomposition techniques . . . 24
2.1.5 Structural equation models (SEM) . . . 25
2.2 Selection of predictors . . . 27
2.2.1 Covariates affected by gender partiality . . . 28
2.2.2 Unobservable constructs . . . 29
2.3 Gender equity at University X: adjusted analyses . . . 30
2.3.1 Hiring . . . 30
2.3.2 Initial rank . . . 30
2.3.3 Initial salary . . . 32
2.3.4 Promotions . . . 32
2.3.5 Annual salary adjustments . . . 34
2.3.6 Departure model . . . 36

3 Estimating higher-level causal effects in complex systems **38**
3.1 Challenges in estimating institution-level effects . . . 41
3.2 Causal inference framework . . . 43
3.3 Simulation-based method of estimation . . . 45
3.3.1 Develop a set of nested models for the observed system behavior 46
3.3.2 Estimate nested models from the observed data . . . 49

3.3.3 Develop a model for the counterfactual system behavior 51
3.3.4 Simulate K realizations of the system and its counterfactual . 53
3.3.5 Estimate average causal effects 59
3.4 Inference for estimated higher-order causal effects 60

4 Assessing sources of information in complex designs 63

4.1 Motivating application : Stepped wedge designs 64
4.1.1 Model and contrasts 66
4.1.1.1 Cross-sectional contrasts 67
4.1.1.2 Longitudinal contrasts 68
4.1.2 Method 69
4.1.3 Simulation study 70
4.2 Assessing sources of information in complex systems 73

5 Simulation study 75

5.1 Data generation under known mechanisms of bias 76
5.1.1 Simulation parameters 77
5.2 Distribution of estimated causal effects 83
5.3 Distribution of estimated variances of estimated causal effects 85
5.4 Contribution of career steps to estimated causal effects 89

6 Gender equity at University X: institution-level analyses 93

7 Discussion **99**

8 Appendix **107**

8.1 Comments on the magnitude and shape of contrast variances (Chapter 4) 107

8.2 Parameters considered to generate stepped wedge data (Chapter 4) . 109

Bibliography **110**

Vita **130**

List of Tables

2.1 University X adjusted analyses: Initial rank 31
2.2 University X adjusted analyses: Initial salary 33
2.3 University X adjusted analyses: Promotions 34
2.4 University X adjusted analyses: Annual salary adjustment 35
2.5 University X adjusted analyses: Departure from the institution 37

4.1 Example of complete stepped wedge design 65
4.2 Variance of $\hat{\beta}_1$: Overall and conditioning on contrasts 71

5.1 Data generation under known mechanisms of bias: Models 81
5.2 Data generation under known mechanisms of bias: Gender coefficients 82
5.3 Average causal effects over 500 simulated datasets by gender partiality scenario 84
5.4 Sources of information of $\overline{RTS}$ 91
5.5 Sources of information of $\overline{HRD}$ 91
5.6 Sources of information of $\overline{TRL2}$ 92
5.7 Sources of information of $\overline{TRL1}$ 92

6.1 University X: Contribution of career models to $\overline{RTC}$ 98

List of Figures

1.1 University X: Gender equity measures 2005-2013 18

3.1 Observed system behavior: University U 49
3.2 Counterfactual system behavior: University U 51
3.3 Academic career simulation procedure 54

4.1 Coefficients λ for cross-sectional and longitudinal contrasts 72

5.1 Distribution of estimated institution-level causal effects by gender partiality scenario 86
5.2 Distribution of estimated variances of institution-level effects for LGP scenario 88

6.1 University X: Example of observed and gender-neutral equity measures 2005-2013 95
6.2 University X: Distribution of estimated causal effects over 100 realizations 96

Chapter 1

Introduction

Gender equity in academia is a topic that has been extensively studied since the early 1920s. At that time, amid the women's rights movement and the ratification of the 19th Amendment, scholars began to assess the presence of women among teaching staff in colleges and universities, and to pay attention to salary, ranks and promotion rate differentials between female and male faculty.

Today however, more than 50 years after the amendment of title VII forbidding sex-based discrimination in US higher education, female faculty are still consistently hired at lower ranks,[2] paid lower salaries[3] and promoted less frequently than men.[4]

Salary, promotions and academic rank constitute components of a complex[5] professional reward system that reflects an institution's needs, priorities and goals. As mechanisms to show value to its faculty members, salary, promotions and academic rank constitute important determinants of productivity and job satisfaction, so assuring that male and female faculty have access to the same opportunities is of the utmost importance.

Many approaches have been proposed to examine the influence of gender on academic rewards, the most commonly used being the Human Capital Theory.[6] Under this theory, gaps in salary, promotion rates and time to promotion are thought to be the result of men and women having different levels of human capital variables (education, experience and productivity), so that if a man has a higher salary than a woman, it must be because he is better educated, more experienced or more productive. However, many studies have shown that, even after accounting for these variables, a portion of the gap remains unexplained.

The vast majority of gender equity research in academia has focused on the individual faculty member as a study unit, using different types of regressions, structural equation models, decomposition techniques, among others, that compare female faculty to otherwise similar male faculty (see Chapter 2). Most of these studies seek to establish an association between gender and academic rewards, usually focusing on only one of the academic reward components. Other studies have looked at causality by staging experiments that prevent the institution from knowing a person's gender during hiring or consideration for promotion.

Most notably, methods are still needed to better assess gender equity at an institutional level, from a causal inference perspective. How do differential gender practices in awarding salaries, promotions and/or academic ranks affect the institution as a whole in terms of its ability to retain faculty members and its desirability as a workplace, as a function of the investment made in its faculty members?

The underlying statistical challenge in addressing these questions corresponds to the estimation of higher-level causal effects in complex systems, i.e., the estimation of causal effects for aggregated units (such as an institution) for which only one observation of the outcomes of interest is available, even though individual-level outcomes are available for the sub-units that make up the system (e.g. faculty members) (see section 3.1). Chapter 3 proposes a simulation based-method, similar to agent-based modeling, to estimate and conduct inference for such higher-level causal effects.

Furthermore, we are interested in investigating how the components of the system affect the system as a whole. Looking at gender equity at the institutional level, this corresponds to assessing the effect of each step of a faculty member's career to the institutional causal effect. What are the sources of institutional gender disparities? How much do differential rates of promotion for male and female faculty contribute to institutional measures of gender partiality? Chapter 4 presents a regression-based method to assess sources of information in complex systems, motivated by the separation of cross-sectional and longitudinal information to estimate the effect of an intervention within a stepped wedge design.

The sensitive nature of gender equity data requires the use of methods that protect the confidentiality of the individual faculty members as well as the institution until results are ready to be published in an aggregated fashion. Therefore, the proposed methods are tested on simulated data with known mechanisms of bias (Chapter 5) and synthetic datasets that mimic information for faculty members of the Johns Hopkins University (JHU) School of Medicine (SOM), between 2005 and 2013, for whom salary, rank and other characteristics are recorded (Chapter 6). At the time of this study, productivity measures were not available to incorporate in the analysis, so the role of productivity in gender equity studies will be discussed but not implemented in this particular application.

In an effort to understand, and address any gender disparities in academic rewards, in 2005 the JHU SOM established a Committee on Faculty Development and Gender which conducts annual cross-sectional analyses monitoring the status of gender disparities in salaries. This thesis aims to extend the efforts of the JHU SOM to achieve a more gender-neutral structure.

Finally, Chapter 7 provides an overview of the proposed methods and its limitations and a summary of the results obtained, along with a discussion of applications other than gender equity, and topics for future research.

The rest of this chapter will be dedicated to setting up the necessary background information that motivates the proposed methods. The term "gender equity in academia" requires special consideration, as each of its components has a very specific meaning. We will describe the academic reward system and provide an overview of the most common outcomes used in gender equity research. Finally, we will describe the data motivating this research and address confidentiality concerns.

1.1 Gender equity in academia

The terms "sex" and "gender", although sometimes used interchangeably, represent different concepts with distinct implications for individuals in terms of health and access to opportunities. Sex refers to the biological characteristics that distinguish men and women (chromosomes, reproductive organs), while gender is a social and cultural construct based on the biological sex differences, including behavior, roles, self-representation, ideology, and psychological traits.[7,8]

Unger and Crawford note the issue with this distinction is not so much due to terminology but to "unresolved conflicts within psychology about the causality of various sex-linked phenomena",[9] so that the term gender is preferred when emphasizing that differences between men are women are a product of social phenomena. From a causal inference perspective, it is important to then distinguish the mechanisms that might produce such differences.

Holland and Rubin coined the maxim "There is no causation without manipulation",[10] defining "causal variables" as those which can be intervened upon. This is the case for example of the study of drug effectiveness, where a patient can be assigned either the drug or a placebo. We then compare the outcome of the patient under the treatment he was actually assigned, to the outcome the patient would have had, had he been assigned to the other treatment, which is the definition of a causal effect.

However, many of the variables of interest in public health are not amenable to causal statements given this definition. This is the case of gender, one of the most used variables in health and social studies. The issue with such attributes is that, since they cannot be manipulated, their causal effects and counterfactual outcomes are not well-defined.[11] This challenge has given rise to extensive discussion in the literature on whether it is even possible to assess causal effects in these cases.[12–15]

One approach to using attributes as causal variables is to shift the perspective of measurement from the subject to the decision-maker, by using instead "perception of the attribute", which can be manipulated.[16] Examples of this can be seen in Bertrand et al. (2003), where names of applicants were changed in resumes to study the influence of race in hiring decisions,[17] and in Goldin et al. (1997) with the use of "blind" auditions to study gender-biased hiring.[18]

We shall then define the casual variable of interest as **gender partiality**, as differential practices by the decision-maker (intentional or otherwise) based on an individual's gender, as opposed to **gender neutrality**, where gender is ignored with respect to opportunities within the institution. Note that this does not mean treating men and women equally, but to ensure they have access to the same opportunities. This concept is known as **gender equity**.[19]

There are situations, however, in which the decision-maker and the individual need to interact face-to-face, making experiments such as the ones mentioned before infeasible. This is the case of the interaction of an academic institution with its faculty members, since individuals need to have some degree of presence at the institution to collaborate on research and/or teach.

Guzmán-Valenzuela and Cortés (2013) define an academic career as the formal mechanism through which individuals move up hierarchical positions over time in an academic institution.[20] These mechanisms reflect the institution's standards and established procedures, as well as its goals and values. This concept is different from *academic trajectory*, which involves the individual's goals and hopes, while an academic career is constrained by institutional practices.

An academic institution, viewed as the collective of its faculty members careers, is a great example of a complex system.[5,21] Institutional policies and actions can be affected by labor market, economic and cultural conditions (contextual effects). Events such as salary increases and promotions, may affect the decisions of faculty members to stay or leave the institution (adaptivity), as well as performance. Performance can, in return and through complex non-linear relationships, determine salary and promotions (feedback loops).

Gender equity in academia has been analyzed from several different perspectives, concentrating mostly on academic rewards (salary, promotion, rank) and female representation at the institution. The percentage of female faculty has generally increased over time, although there exist marked differences across fields.[22] From 1921, when 12.0% of all teaching positions and 4.0% of full professorships in coeducational colleges and universities nationwide were held by women,[23] with no woman holding any grade of professorship in colleges for men only, 24.3% of full-time faculty are women in 2013.[24]

1.1.1 Representation

Studies focusing on representation attempt to explain unequal sex ratios at hiring, departure, or at some fixed point in time in terms of human capital variables, productivity, academic climate, and satisfaction.

The analysis of representation at hiring focuses on three questions: are men and women equally likely to: apply for an academic position, be made an offer of employment, and to accept an offer if one is made.

The composition of the applicant pool reflects both the practices and values of the institution in the search and recruitment process, and market, social, cultural or economic conditions in the pool of applicants. It follows that unequal gender ratios may be outside the control of the institution.[25] The decision to apply is determined by characteristics of the institution and characteristics of the applicant,[1] so without records of actual applicants it might prove problematic to ascertain gender partiality using national figures (such as the number of women with doctorates in the field) and figures of new graduates would be relevant only to entry-level or assistant professor rank. Furthermore, the candidate selection process "can be difficult to quantify",[1] since qualitative personality traits and fit to the institutions are important factors that influence offers.

According to the National Science Foundation, 46.1% of doctorate recipients in 2014 were female, although this percentage is lower for the physical sciences and engineering (28.7% and 22.9% respectively).[26] Also for 2014, the Association of American

Medical Colleges reports that 47.6% of medical graduates are female.[27] However, the percent of women who choose to pursue an academic career might be lower than the fraction of men that do,[1,28] but this varies largely by field and type of institution. Out of the women who do apply, they are more likely than men to receive an offer.[28]

Some literature suggests that there might be bias in the hiring process, making it more difficult for women with the same qualifications as men to be hired.[1,29,30] However, Ceci and Williams (2010) report that in the case of math-intensive fields, these disparities are due to "differences in resources, abilities and choice" instead of discrimination,[31] and hiring experiments have also shown a preference to hire female faculty members in STEM.[32]

Female faculty attrition has been shown to be higher than male and to be linked to job dissatisfaction, related to lack of opportunities, mentorship, advancement opportunities, and generally an unattractive work climate.[1,33–35]

1.1.2 Academic rewards

Studies focusing on academic rewards attempt to explain differences in salary and rank as a function of institutional and individual characteristics, particularly referring to human capital and performance measures (e.g. number of publications, grants accepted, number of citations). Gender disparities usually manifest in one of two ways: a different reward structure for female and male faculty (the return on covariates is different by gender) or the same reward structure is present but a constant factor for salary or rank is subtracted for women.[36]

The most prevalent research in this area are salary equity studies, with outcomes such as salary-at-hire, changes in salary over time, or salary at a fixed point in time.[3,6,37] Ranks have been studied in terms of rank-at-hire, at a fixed point or over time, as well as probability of promotion, time to promotion and probability to achieve tenured status, among other outcomes.[38–40]

Equity studies became popular in the late 1960s, following two important changes in legislation. The Equal Pay Act of 1963 aimed at eliminating differences in wages on the basis of sex, requiring equal pay for jobs performed under similar conditions and that required the same level of "skill, effort and responsibility" (except in cases where wages are determined by seniority, merit, quantity or quality of production, or when wage differentials are based on factors other than sex).[41] Title VII of the Civil Rights Act of 1964 further extended legislation by forbidding discrimination on the basis of sex, race, religion or national origin, requiring that jobs that are of the same value for the employer, even if they are very different jobs, should be compensated equally.[42] Title VII was further amended in 1972 to explicitly forbid discrimination in academic employment.[6]

Even with these and newer changes in legislation[1], research conducted over the past 50 years suggests a differential favoring male faculty members, after accounting for relevant covariates, that persists to this day.

[1]Such as the Lilly Ledbetter Fair Pay Act of 2009[43]

Studies using national datasets show that the earnings differences between male and female faculty in science and engineering has decreased over time,[44] with estimated statistically-significant differences between -20.0% and -12.0% in the 1960s, and -6.0% to -10.0% in the 1980s (with some studies showing no significant differences), while recent single university studies show gaps lower than 5.0%. The University of Minnesota found in 2011 that their male faculty were paid 2.2% more[45] than their female faculty[2]. At the University of California-Berkeley, 2015 data show that women earn between 1.8% and 4.3% less than otherwise similar men.[46] The College of Liberal Arts at the University of Austin - Texas identified that salaries for female faculty were between 2.0% (full-level) and 6.0% (assistant-level) lower than male's (2013).[47] The University of Indiana-Purdue found a 3.0% percent gap in average salaries in 2014, which has remained since 1998.[48]

Women in academic science and engineering have also been shown to be historically underrepresented in higher ranks and to have a lower chance of being promoted than otherwise similar men in studies using national datasets.[44] This trend has been present since the 1960s with varying degrees of severity. Ash et al. (2004) and Jena et al. (2014) show this trend also exists for women in academic medicine, with 66.0% of men but only 47.0% of women with 15-19 years of seniority holding full professorships in 1995-1996,[49] and 11.9% versus 28.6% of full professor ranks going to women in 2014.[50]

[2] Adjusting for variables thought to influence salary, with the exception of merit-related covariates, which were not available.

1.2 Data and confidentiality

This study was motivated by annual gender equity reviews between 2005-2013 conducted by the Johns Hopkins Biostatistics Center on behalf of the JHU SOM Committee On Faculty Development and Gender. Faculty included in the analyses are full-time faculty in the ranks of Assistant Professor through Full Professor, excluding deans, department and institute directors and faculty who were previously in these leadership positions. Note that not all faculty members are followed for the same period of time, reflecting the dynamic nature of academic employment.

The annual dataset contains information of salary, rank, and other demographic characteristics of the faculty members including age, education and department. Information is obtained from the SAP Enterprise Management System and validated by the JHU SOM Registrar's Office. Salary data corresponds to full-time equivalent (FTE) salary, which is comprised of the base salary (part A) plus any supplemental salary for administrative, educational or clinical roles assumed by the faculty (part B) (FTE=A+B). The dataset also contains information on total salary, which adds any bonuses (part C) defined in the individual departmental compensation plans or that have been agreed upon by the faculty and department (Total salary=A+B+C).

This information, however, is necessarily **confidential**. Gender equity studies use extremely sensitive information that, if made public could potentially harm faculty or the institution, even if it is committed to gender equity in hiring, promotion, and salary setting policies.

In studies that use national datasets in which information from multiple universities is available, the name and location of the universities are suppressed to retain confidentiality. Also, the use of multiple universities in an analysis, prevents identification of an individual faculty member and his/her personal characteristics or conditions of employment.

However, for single university studies, care must be taken to protect both the individual faculty members' identity and the institution as a whole. This situation has been addressed in the literature by adding noise to the data, excluding outliers or working with aggregated data to avoid identification of individual faculty members. However, it has been noted that these techniques may not fully protect anonymity[51] and may compromise data usability and conclusion validity.[52] One approach to overcome these limitations is the use of multiply imputed synthetic data: datasets generated by sampling from a hypothesized data generating mechanism, so that a set of characteristics of the original data are preserved, e.g. moments of the joint distribution. Inferences are obtained by combining results from all datasets.[53]

For this thesis, a single synthetic dataset was generated from the JHU SOM data, and **this dataset was used in place of the institutions' original data**. Note that in this case only a single dataset is needed, since the goal is neither to estimate population parameters nor to make inferences for the original data, but to have a dataset with which to test methods while preserving the confidentiality of faculty members and the institution as a whole.

The synthetic dataset was constructed in two stages:

1. Generation of the initial characteristics of faculty members.

   First, characteristics of both existing faculty members at the institution at the beginning of the study (year 2005) and newly hired faculty members for each subsequent year in the study (years 2005 to 2013) were generated.

   The number of unique faculty members, i.e., distinct faculty members who are present at the institution at least one year, was chosen to be similar to the number of unique faculty members at the JHU SOM in the period 2005-2013. Each one of these unique faculty members is assigned a year of entry to the dataset by sampling with replacement from the distribution of JHU SOM.

   The rest of the faculty characteristics at entry are generated using non-parametric "simple synthesis".[54] This procedure uses regression and classification trees to predict the faculty characteristic of interest as a function of chosen covariates using a prediction algorithm that is driven by the data and makes no parametric assumptions.

   Each variable in the dataset is then generated as a prediction from a model that conditions on the variables that precede it in the dataset. The faculty characteristics are generated sequentially as follows: department, gender, race, age, degree, arrival[3], rank, time in rank and FTE salary. This means that, for example, gender is generated conditioning on year, department, and existing/new

[3]Whether the faculty member is a new hire or not

hire status. This process continues until the generation of FTE salary, which is based on all the preceding faculty characteristics.

2. Generation of the longitudinal academic careers.

   The longitudinal academic careers were synthesized following the logic presented in Figure 3.3. Conditional on initial characteristics, we used the parametric career models defined in section 3.3.2 to simulate the promotion, change in salary and departure processes.

For the rest of this work, we shall refer to the generated synthetic data as data for University X.

1.3 Gender equity at University X: descriptive analyses

University X had a total of 2458 unique faculty members with information for at least one year in the period 2005-2013, with the size of the institution increasing from 1362 faculty members in 2005 to 1654 in 2013 (Figure 1.1a). University X hires on average 130 new faculty members each year (standard deviation SD=25), and 85 faculty members depart on average each year (SD=20).

By 2013, University X employed 38.0% women (Figure 1.1a) and 7.5% underrepresented minorities, an increase from 32.1% and 5.7% respectively in 2005, and with a faster increase for female underrepresented minorities (9.3% versus 6.4% for males in 2013). The percentage of new hires who are women at University X averages 41.4% with a peak of 47.0% in 2008.

The majority of faculty members (92.2%) stay at least 10 years at the institution, with very similar time distributions for female and male faculty. The number of women leaving University X has increased over time, with approximately 40.0% of departures being female faculty members in the period 2010-2013.

The difference in median age for female as compared to male faculty varied from between 2.2 years (2006) and 4.6 years (2011). For 2013, the median ages were 48.0 and 43.8 years for male and female faculty members respectively.

Approximately 73.0% of faculty members have MD degrees, while 13.5% have PhD degrees. This distribution has not changed over time, with slightly lower proportions for female faculty with an MD degree (70.0%) and declining number of female PhD holders (13.4% in 2005 versus 10.2% in 2013).

Ranks at University X have remained relatively stable over time at the ratio 2:1:1 for Assistant, Associate and Full Professors (Figure 1.1b). This distribution however is different by gender. For 2013, 57.4% of female faculty were at the Assistant rank, compared to 46.1% among male faculty. As rank increases, the percentage of female faculty is smaller: for 2013, 43.3%, 36.9% and 28.1% of Assistant, Associate and Full rank positions respectively were held by women.

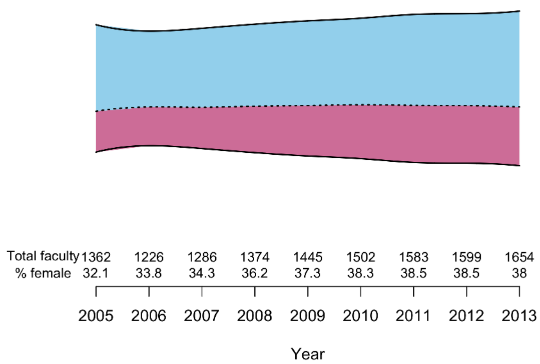
A total of 442 faculty members were promoted in 2005-2013, 35 of whom were promoted more than once. 65.4% of the faculty members promoted once and 80.0% of the faculty members promoted twice were male. Median time in rank is similar for female and male faculty in Assistant and Associate ranks, but male faculty have longer time at Full professor rank (11.4 years versus 8.9 for females).

On average, 90.0% of newly hired faculty members are hired at the Assistant Professor level. Average initial salaries for new Assistant Professors at University X have increased from approximately 107,600$ (SD=27,613$) in 2005 to 141,000$ (SD=33,395$) in 2013, although this varies by department, with the average initial salary by department ranging from 93,400$ in 2013 (SD=23,276$) to 163,500$ (SD=83,595$).

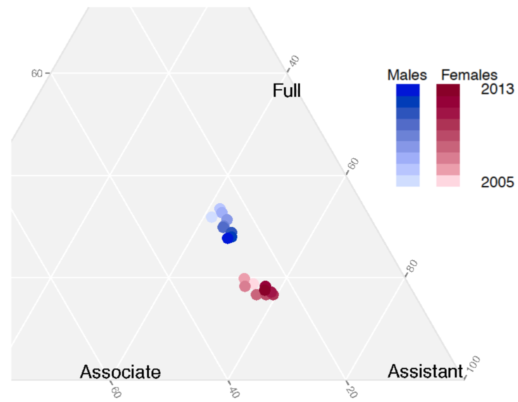
Initial average salary for women hired as Assistant Professors has been historically lower than males, although this difference is not statistically significant in unadjusted analyses except for 2007 and 2013, when female initial salaries were on average 20,329$ ($p=0.0016$) and 11,806$ ($p=0.0542$) lower than males, respectively.

Over time, women experience 0.3% higher average annual increases in salary than men ($p<0.001$) (Figure 1.1c).

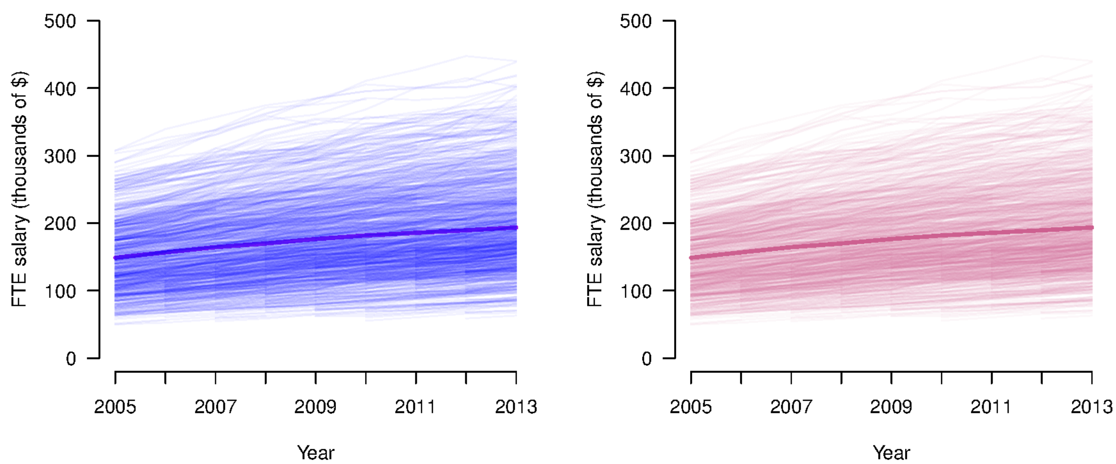
All of these results represent unadjusted analyses of individual-level gender equity measures. Confounding variables such as human capital and performance/productivity were not taken into account. Chapter 2 will deal with existing statistical techniques to assess gender equity at individual level.



(a) Distribution of faculty by gender over time



(b) Distribution of faculty ranks by gender over time. Axes represent % faculty in each rank. For example, among female faculty in 2013, 18.0% were full professors, 25.0% associate professors and 57.0% assistant professors.



(c) Individual FTE salary by gender over time

Figure 1.1: University X: gender equity measures 2005-2013

Chapter 2

Literature review: statistical approaches to asses gender equity

Despite the enormous capabilities of existing statistical approaches to assess gender equity, these methods are insufficient to investigate equity at the institutional level for single-institution studies, from a causal inference perspective. Section 3.1 will elaborate on the specific challenges of shifting the perspective of measurement from the individual to the institution, while section 3.3 will propose a method that extends existing approaches to assess gender equity at institution level.

This section then focuses on describing the most common statistical methods that have been used to assess gender equity at the individual level in the literature, i.e., to assess the association between gender and academic rewards outcomes measured **for each faculty member**. Then, individual faculty members are used as replications of the perception and action of the institution on a person's gender, and these results are used to reach conclusions at the institution level.

In terms of study design, literature on equity studies varies widely as a function of the number of universities included in the study (single or multiple), the number of disciplines included (single or multiple), the type of design (observational or experimental), and the number and nature of time points to analyze (cross-sectional, repeated independent samples or longitudinal data).

The 1970s were characterized by cross-sectional, single university studies that aimed at evaluating salary systems, determining the methodology to assess inequities, monitor changes in policies and implement actions regarding sex-discrimination claims.[55] With the availability of national datasets, research in the 1980s focused on multi-university studies using repeated cross-sectional samples over time, aimed at studying trends and generalizing results from case studies, as well as assessing the effects of anti-discrimination measures put into effect in the previous decade.[55]

This trend toward multi-university studies over time continued into the 1990s, but more studies started to focus on methodological issues, such as the inclusion of gender-biased covariates in regression models.[55] Study designs and analytic techniques evolved to reflect the need for more sophisticated analysis, that could take into account the multidimensionality of the problem.

The most widely used statistical techniques to assess gender equity are regression models, but other techniques have also been used. Earlier studies were mostly unadjusted comparisons of means and percentages that were likely confounded by variables known to be determinants of academic rewards and representation, such as level of experience and productivity. Section 2.1 presents an overview, by no means exhaustive, of methods used in the gender equity literature that allow to control for potential confounders.

We finish this chapter with an adjusted analysis of the gender equity situation at University X (section 2.3).

2.1 Existing methods

2.1.1 Matching

Also known as counterparting,[6,56,57] matching involves paired comparisons of female faculty to male faculty with the same or similar key characteristics such as rank and measures of productivity. This technique was commonly used in the 1970s and 1980s when the number of female faculty members was smaller. Sets of matches were sent to a review committee to evaluate academic rewards and determine if an adjustment should be made.

This approach is largely qualitative because it relies on opinions of a committee for evaluation making it prone to gender biases. It carries the additional problem that matches might be difficult to find in high dimensional space. There is also the risk of over-matching,[57] that is, matching on variables in the causal pathway of gender bias.

2.1.2 Compa-ratios

Bereman and Scott (1991) suggested the use of compa-ratios[58] in single-university studies as an alternative to regression techniques, to avoid issues of interpretability while still being able to adjust for rank, discipline, productivity and time-in-rank. Compa-ratios are obtained by dividing individual salaries by the salary midpoint for faculty in the same "pay grade" (faculty with the same rank and field for example), and the ones that fall below a certain threshold are reviewed for potential causes of low pay (such as productivity).[6] This technique requires the assumption that rank and academic field are not affected by gender biases.

2.1.3 Multivariable regression

Multivariable regression in any of its forms (linear[1], logistic,[22,39,59] probit,[60,61] proportional hazards models,[38,40,62] random effects models,[63–65] among others) is the method of choice for analyzing gender equity in representation and academic rewards. The popularity of these techniques is based on the fact that regression models can accommodate different natures of outcomes and predictors, and they allow the estimation of differences between female and male faculty while adjusting for potential confounders.

[1] See Loeb and Ferber (1971), Katz (1973), Farber (1977), Hirsch and Leppel (1981), Tolbert (1986), Pfeffer and Ross (1990), Formby (1993), Ginther (1999), Monks and Robinson (2000), Toutkoushian et al. (2007), and Binder et al. (2010), just as a few examples.

The simplest regression approach is the single-equation method, in which a single regression model is fitted to the all the data, including both female and male faculty members. In this case, gender is included in the model as a dummy variable and its coefficient reflects the "unexplained" differences between females and males after adjusting for the rest of the variables in the model. This approach assumes that the gender effects on salary or promotion are the same for different values of the covariates, i.e., the reward structure is the same for female and male faculty.

The assumption of equal reward structure may not hold, however. Female and male faculty might be rewarded differently as a function of their characteristics. For example, while a male faculty member might get an increase in salary of 500$ for each additional publication, a woman might get 250$ only. This can be incorporated into the model by adding interaction terms with gender[74] or by running separate regressions by gender (multiple-equation method).[75] The coefficients for the female and male models can then be compared using Chow's test[60,70] or differences in means tests.[76,77] This approach has the problem that each model is run on a smaller sample, so these estimates are less reliable than the single-equation model approach.

Also popular is the use of "reverse regression", introduced by Conway and Roberts as a complement to regular regression.[78] In ordinary regression, we assess whether the outcome, compensation for example, is the same for otherwise similar female and male faculty in terms of qualifications. The goal of reverse regression is to assess whether qualifications differ among equally paid female and male faculty.[79,80] An example of this technique is when logistic regression is used to predict a faculty member's gender in terms of demographics and academic rewards.[81] Reverse regression might yield very different qualitative results when compared to regular regression, and it has been criticized as not yielding unbiased estimates of gender partiality.[82]

Several issues related to the selection of predictors and construction of the regression model apply to gender equity models. The issues of multicollinearity, inclusion of irrelevant variables, omission of relevant variables, incorrect specification of the functional relation between the response and the predictor, outliers and influential observations, non-normal distribution of residuals, non-constant variance and correlated errors should all be assessed through diagnostic tools and addressed to get a well-specified model. There are two issues, however, that are of particular importance when conducting gender equity analysis: the inclusion of covariates potentially affected by gender (e.g. academic rank) and the inclusion of unobservable constructs, specifically, faculty productivity. We will discuss these issues in section 2.2.

2.1.4 Decomposition techniques

Decomposition techniques are based on the multiple-equation regression approach[2]. After fitting separate regression models for female and male faculty members, the observed differences between the two are further decomposed into components by comparison to a gender-neutral scenario: one component attributable to differences

[2] Examples of studies using these techniques can be seen in Buzan and Hunt (1976), Barbezat (1987), Haberfeld and Shenhav (1990), and Ashram (1996).

between female and male faculty in terms of covariates (education, experience, productivity, among others), and one or more components of unexplained differences, which are commonly used as a measure of discrimination. This approach was introduced simultaneously in 1973 by Oaxaca and Blinder to assess gender-based discrimination in salaries[84–87] and later extended to logit and probit models.[88,89]

Two-equation models assume that gender partiality manifests as either female or male faculty being under or overcompensated, but not both at the same time.[75] The decomposition is achieved by assuming the gender-neutral structure corresponds to one of the genders.[84] By comparing the results of the models using the female and male structures we can get an interval for the estimated "discrimination" coefficient.

Three-equation models allow for gender partiality to manifest as both overcompensation for men and under compensation for women, at the same time.[90] In this regard, Neumark suggested the use of the coefficients of a pooled model (excluding gender) as the gender-neutral structure, while Cotton (1988) suggested using the weighted average of the male and female coefficients to get the gender-neutral structure.[91]

2.1.5 Structural equation models (SEM)

These techniques are used to asses complex relationships between the academic reward and representation outcomes and faculty characteristics, by positing directionality and, potentially, temporality between multiple outcomes and covariates. They also allow for the incorporation of latent variables (such as true productivity), by hypothesizing causal pathways between measured and latent covariates.

Unlike multivariable regression, which usually focuses on a single outcome, SEMs allow the assessment of the interdependence among multiple outcomes, and to investigate both direct and indirect effects of measured covariates and latent constructs using a system of equations. Furthermore, while traditional regression techniques assume no measurement error, structural equation modeling can accommodate imperfect measures.

These characteristics make SEMs particularly suited to analyze gender equity in academia, since rank and salary are closely related. For example, faculty in higher ranks tend to have higher salaries and promotions are usually accompanied by a boost in compensation. Rank can also be potentially affected by salary, in that a faculty member that deserves a pay raise might get promoted to allow for the raise if the university salary policies include salary ceilings by rank. Departures from the institution might be based on salary and promotion decisions. And all of these aspects might be affected by gender.

Although these techniques have been used to assess the gender wage gap in the labor market, there are few applications specific to academia. Smart (1991) used "causal modeling" to assess direct and indirect gender influences on promotions and salary.[97] Fisher, Motowildo and Werner (1993) used correlation and path analysis to look for evidence of whether gender is a determining factor of academic rank and salary.[98] More recently, the University of California at Davis (2013) used structural equation modeling to identify "specific processes and historical trends that have the greatest influence on salary disparities, and thus are potential targets for policy modification", gender being one of the variables in their study.[99]

Other techniques such as the Salary Kit method,[100] discriminant analysis,[60,101] classification models,[102] and biplots[103] have also been used, but they are not as popular as single- or multiple-equation regression models or decomposition techniques.

2.2 Selection of predictors

The choice of variables to include in any model should be grounded in the conceptual framework for the process being studied. In general, gender equity studies in academia include approximately the same concepts (even if sometimes operationalized differently) of human capital, structural and market considerations.[6] This, evidently, depends on the data available for each particular study. Some of the most frequently used covariates are:

- Education: highest degree earned, institution of degree, academic field.
- Experience: time since degree, years at the institution, years in rank, age.
- Performance: number of publications, publication journal impact, books, number of citations, number and amount of grants, student and peer evaluations, dissertations supervised, hours spent on teaching and research.

- Service: service on administrative committees, leadership positions.
- Institution: type, size, location, prestige, unionization, discipline.
- Labor market: average national salary, unemployment rates, number of graduates in the field.

There are two particularly controversial points regarding the selection of predictors in gender equity analyses: the inclusion covariates potentially affected by gender (e.g. academic rank) and the inclusion of unobservable constructs (e.g. faculty productivity).

2.2.1 Covariates affected by gender partiality

Many of the covariates usually used to assess differences in faculty rewards and representation in academia might be affected by gender partiality. This is the case of productivity measures,[104] merit-based variables and academic rank. Unfriendly work environments, lack of mentorship, bias in awarding academic honors, are means by which gender partiality influences "productivity". Including such variables in a model will lead to underestimating the differences between female and male faculty, because such analyses assume that men and women have had the same opportunity to "acquire the attributes" that affect the academic reward outcomes.[6,93]

However, failing to include all relevant covariates in the model will also lead to biases due to omitted variables. Rank, for example, has been shown to be one of the most important predictors of salary,[25,105] as initial salary is based on rank and promotions are accompanied by salary increases. For example, Boudreau (1997) showed, using both hypothetical and actual data free of gender bias, that failing to include rank in the model produces results that indicate the presence of gender bias.[106]

Two commonly used approaches are to model the covariates potentially affected by gender first, and only include them in the model for the outcome of interest if there is no evidence of gender bias,[59,61,98,104,107] or to run separate models with and without including these covariates and presenting results for both.[37,60,108,109]

2.2.2 Unobservable constructs

The main example of an unobservable construct in gender equity studies in academia is faculty productivity, and the issue with such variables is that they can only be imperfectly measured through observable indicators such as number of publications or citations.[110] Furthermore, the observable indicators might represent measures of quantity (and not quality) of work.[111]

Given this issue, some authors choose to exclude productivity measures from their models,[104] sometimes citing that men and women are equally productive[112,113] so this construct, if measured without error, would not confound the relationship between academic rewards and gender. However, we can only observe if gender is related to observed measures of productivity and, if gender is related to true productivity, failing to include it in the model will cause bias due to omission of a relevant variable.

2.3 Gender equity at University X: adjusted analyses

2.3.1 Hiring

University X does not record faculty applicant characteristics, so assessing gender equity in the hiring process is not possible. However, we can compare characteristics of the male and female faculty at the time of hire.

Using a logistic regression model for gender (female vs. male), we find that newly hired older faculty members and PhD holders are less likely to be female, with OR=0.95 (95% CI=[0.93,0.97]) for age and OR=0.60 (95% CI=[0.41,0.88]) for PhD vs. MD, respectively, after adjusting for year, rank and department of hiring.

2.3.2 Initial rank

Initial rank was studied using an adjacent category logit model with non-proportional odds for gender, adjusting for degree, age of hire (quadratic function), and department[3](Table 2.1).

Newly hired female faculty have 42.0% lower odds of holding an initial rank of Associate Professor compared to Assistant Professor, but 3 times the odds of being hired as a Full Professor compared to Associate when comparing to male faculty.

[3]As measures of education, experience and field

	Pooled		Female		Male	
	$\hat{\beta}^{*}$	95% CI	$\hat{\beta}^{*}$	95% CI	$\hat{\beta}^{*}$	95% CI
Female						
Full vs Associate	1.16	(0.36,1.96)				
Associate vs Assistant	-0.54	(-1.05,-0.03)				
Degree [a]						
MD-PhD	0.20	(-0.20,0.60)	-0.14	(-0.90,0.62)	0.39	(-0.11,0.89)
PhD	-0.02	(-0.38,0.34)	-0.30	(-0.9,0.30)	0.19	(-0.28,0.67)
Other	0.49	(0.12,0.87)	0.65	(0.15,1.15)	0.34	(-0.24,0.91)
Age at hire [b]						
Linear	-0.014	(-0.04,0.02)	-0.034	(-0.09,0.02)	-0.003	(-0.04,0.03)
Quadratic	0.002	(0.000,0.003)	0.003	(-0.002,0.008)	0.001	(-0.003,0.006)

[*] $\hat{\beta}$ represents difference in log odds of being hired at a higher rank comparing to the reference category

[a] Reference category is MD degree

[b] Centered at mean

Table 2.1: Log odds of being hired at a higher rank: model coefficients and 95% confidence intervals (CI). Adjacent category logit model with non-proportional odds for gender. Negative coefficients indicate lower log odds of holding a higher rank (Full vs Associate or Associate vs Assistant).

2.3.3 Initial salary

Initial salaries were assessed using a heteroskedastic mixed effects model for log FTE salary with a random intercept at department level and different residual variances by department.

Initial salaries are 6.0% lower for females than otherwise similar males (Table 2.2). Initial salary is lower for MD-PhDs and other degrees compared to MDs, and lower for Associate and Assistant Professors compared to Full Professors. Adjusting for rank has no effect on the gender coefficient.

The effect of degree and rank on initial salary is significantly different for female and male faculty ($p<0.001$). Female MD-PhD holders have initial salaries that are 20.5% lower than female MDs, while male MD-PhD holders make 12.2% less than male MDs. Differences in initial salary by rank are more accentuated among male faculty, for whom an Associate professorship implies a 28.1% lower salary compared to Full Professor, instead of the 21.3% lower salary comparing female Associate Professors to female Full Professors. The same holds when comparing Assistant and Full professors, with females having a 40.5% lower salary compared to 47.8% lower for males.

2.3.4 Promotions

Promotions were studied using ungrouped Poisson regression, i.e., log-linear models for being promoted which use single units of faculty-time.[114] These models yield

	Pooled		Female		Male	
	$\hat{\beta}^*$	95% CI	$\hat{\beta}^*$	95% CI	$\hat{\beta}^*$	95% CI
Female	-0.06	(-0.09,-0.04)				
Degree [a]						
MD-PhD	-0.17	(-0.20,-0.13)	-0.23	(-0.29,-0.17)	-0.13	(-0.18,-0.09)
PhD	-0.43	(-0.46,-0.40)	-0.38	(-0.42,-0.34)	-0.48	(-0.52,-0.44)
Other	-0.12	(-0.16,-0.08)	-0.15	(-0.21,-0.09)	-0.1	(-0.15,-0.05)
Year [b]	0.03	(0.02,0.03)	0.03	(0.02,0.03)	0.03	(0.02,0.04)
Rank [c]						
Associate	-0.27	(-0.34,-0.19)	-0.24	(-0.34,-0.13)	-0.33	(-0.43,-0.23)
Assistant	-0.58	(-0.65,-0.52)	-0.52	(-0.60,-0.44)	-0.65	(-0.74,-0.56)

[*] $\hat{\beta}$ corresponds to mean difference in log initial salary comparing to the reference category

[a] Reference category is MD degree

[b] Centered at 2005

[c] Reference category is Full Professor

Table 2.2: Log FTE initial salary: model coefficients and 95% confidence intervals (CI). Heteroskedastic mixed effects model for log FTE salary (random intercept at department level and different residual variances by department) fit with REML. Negative coefficients indicate relative lower geometric mean of initial salary.

results equivalent to proportional hazards regression. There are not enough promotions to run separate regressions by gender.

The hazard of promotion to Associate Professor is 18.1% lower for female compared to otherwise similar male faculty, but there are no gender differences in the hazard of promotion to Full Professor (Table 2.3). Older faculty are less likely to be promoted. MD-PhD and PhD holders are more likely than MDs to be promoted to Associate Professors, while only MD-PhDs are more likely than MDs to be promoted to Full Professors. The hazard of promotion increases with time in rank up to 9.5

years for Assistant professors and 10.9 years for Associate Professors, after which the hazard of being promoted declines over time.

	Assistant to Associate		Associate to Full	
	$\hat{\beta}^*$	95% CI	$\hat{\beta}^*$	95% CI
Female	-0.20	(-0.42,0.02)	-0.18	(-0.54,0.17)
Time in rank [a]				
Linear	0.34	(0.29,0.39)	0.43	(0.34,0.53)
Quadratic	-0.05	(-0.06,-0.04)	-0.04	(-0.06,-0.03)
Degree [b]				
MD-PhD	0.55	(0.25,0.83)	0.96	(0.52,1.39)
PhD	0.33	(0.06,0.59)	0.55	(0.12,0.96)
Other	-0.01	(-0.54,0.46)	0.40	(-0.34,1.02)
Age [c]	-0.02	(-0.04,0.00)	-0.11	(-0.15,-0.07)

* $\hat{\beta}$ corresponds to the difference in log hazard of promotion comparing to the reference category

[a] Centered at 6 years

[b] Reference category is MD

[c] Centered at mean

Table 2.3: Hazard of promotion: model coefficients and 95% confidence intervals (CI) for Poisson promotion models. Negative coefficients denote lower hazard of being promoted. Models are specific to a particular rank.

2.3.5 Annual salary adjustments

Annual salary adjustments were modeled using the log of the change in salary comparing the current salary to the previous year salary as the outcome in a mixed effects regression model with random intercepts for faculty and department, to account for clustering of faculty in department and correlation in changes in salary within a faculty member over time.

	Pooled		Female		Male	
	$\hat{\beta}^*$	95% CI	$\hat{\beta}^*$	95% CI	$\hat{\beta}^*$	95% CI
Female	0.002	(0.001,0.003)				
Time in rank ^{a}	-0.007	(-0.008,-0.006)	-0.007	(-0.009,-0.005)	-0.007	(-0.008,-0.006)
Rank and Promotion ^{b}						
Associate not promoted	0.0004	(-0.0009,0.0017)	-0.0004	(-0.0027,0.0018)	0.001	(-0.0007,0.0027)
Associate just promoted	0.097	(0.094,0.100)	0.094	(0.089,0.099)	0.099	(0.096,0.103)
Full not promoted	0.004	(0.002,0.005)	0.002	(-0.001,0.005)	0.005	(0.003,0.007)
Full just promoted	0.076	(0.071,0.081)	0.059	(0.050,0.068)	0.084	(0.078,0.090)
Previous salary ^{c}	-0.016	(-0.018,-0.015)	-0.016	(-0.019,-0.013)	-0.017	(-0.019,-0.015)
Year ^{d}						
Linear	-0.005	(-0.005,-0.004)	-0.005	(-0.005,-0.004)	-0.005	(-0.005,-0.004)
Spline (knot=10 years)	0.003	(0.002,0.004)	0.003	(0.002,0.004)	0.003	(0.002,0.004)

* $\hat{\beta}$ corresponds to the difference in mean log relative change in salary compared to the reference category

^{a} Centered at 6 years, coefficient indicates changes per 10 years in rank

^{b} Reference category is Assistant not promoted

^{c} Centered at mean

^{d} Centered at 2005

Table 2.4: Annual salary adjustment : model coefficients and 95% confidence intervals (CI) for mixed effects regression model for relative change in log FTE salary: $\log \dfrac{FTE\ salary(t)}{FTE\ salary(t-1)}$, with random intercepts at person and department level. Negative coefficients indicate a relative decrease in FTE salary.

Annual increases in salary are 0.2% higher for females than for their male counterparts, after adjusting for time in rank, rank and promotion, previous salary, and calendar year (Table 2.4). Larger annual increases in salary are associated with promotions, while longer time in rank and higher previous salary are associated with smaller changes in salary. Annual increases in salary have been decreasing from 2005 to 2010, but remain roughly constant from 2010 to 2013.

Female faculty get smaller annual increases in salary than their male counterparts. Looking at faculty at Full Professor rank, promoted female faculty receive a salary adjustment that is 5.7% higher than females not promoted. However, promoted males receive a 7.9% higher salary adjustment than males not promoted. Female faculty recently promoted to Full Professor then receive a 2.2% significantly lower adjustment than males just promoted to Full Professor ($p<0.001$). There are no differences in the salary adjustments when a faculty is promoted to Associate Professor comparing female and male faculty.

2.3.6 Departure model

Using using ungrouped Poisson regression, we found that the hazard of departure from University X is not associated with gender. Full Professors are more more likely to leave University X than both Assistant and Associate professors, as are faculty with negative changes in salary with respect to the previous year. Faculty with higher log FTE salary and higher salary increases are also less likely to leave.

	$\hat{\beta}^*$	95% CI
Female	-0.10	(-0.29,0.09)
Change in salary [a]		
<0%	0.14	(-0.44,0.72)
(0,1]%	-1.14	(-1.85,-0.48)
(1,2]%	-0.90	(-1.45,-0.35)
>2%	-0.65	(-1.05,-0.2)
Rank [b]		
Associate	0.35	(0.08,0.63)
Assistant	0.70	(0.42,0.98)
Log FTE salary [c]	-0.03	(-0.33,0.26)
Time in rank [d]	0.03	(0.01,0.04)

* $\hat{\beta}$ corresponds to differences in log hazard with reference category

[a] Reference for Change in salary is Assistant not promoted

[b] Reference category for Rank is Full professor

[c] Time in rank centered at 6 years

[d] Log FTE salary centered at mean

Table 2.5: Hazard of departure: model coefficients and 95% confidence intervals (CI) for Poisson regression model. Negative coefficients denote lower hazard of departure.

Chapter 3

Estimating higher-level causal effects in complex systems

The study of causal effects is ubiquitous in the research literature. The Rubin Causal Model[12,119,120] has provided a structured way of estimating causal effects based on the potential outcomes framework.

Suppose we are interested in the causal effect of a binary exposure Z on an outcome Y for a target population. We define the individual causal effect for unit i as the comparison of the potential outcomes $Y_i(Z = 1)$ versus $Y_i(Z = 0)$, i.e, the outcome unit i would have, had it been exposed to Z, versus the outcome unit i would have, had it not been exposed to Z. In the case of gender equity in academia, the exposure is “gender partiality”. An example of a causal effect would be the comparison of the salary faculty member i would have had at Institution X, had Institution X been gender partial ($Y_i(Z = 1)$) versus the salary faculty member i would have had, had Institution X been gender neutral ($Y_i(Z = 0)$).

We are generally interested in the average[1] of the individual causal effects for the target population[2]. The average treatment effect (ATE) and the causal risk ratio are examples of such causal effects. In practice, only one of the individual potential outcomes is observed, so individual causal effects cannot be directly calculated. Population causal effects, on the other hand, can be estimated but require assumptions to be made about the potential outcomes.[120]

The vast majority of causal inference research focuses on the individual as the level of analysis. There are situations however in which the outcome of interest is measured at a higher level, as a function of a cluster of units, and where all units in the cluster are either exposed or unexposed to Z. An example of this would be to assess the causal effect of gender partiality on the total amount in dollars paid in salary for a given department within a university. Departmental policies can be either gender partial or gender neutral. In this case, the unit of analysis is the department. With some required assumptions,[120] we can estimate the effect of gender partiality by comparing the outcomes of the gender partial departments to the outcomes of the gender-neutral departments. It is likely, however, that gender partial policies at the university level permeate departmental policies, and so we would be interested to assess the effect of gender partiality at the university level.

[1] Or any other aggregating function appropriate to the nature of the outcome

[2] This is what Hernan calls "aggregated" or "population" causal effects.[11]

We consider the case where the level of measurement is the collective of all units, e.g., the university as the collective of all its faculty members. We are interested in quantifying the causal effect of gender partiality, a university-level phenomena, on outcomes also defined at the university-level, such as the total amount in dollars paid to faculty over a period of time. Since we are interested in assessing gender equity for one particular university, this is akin to estimating an **individual causal effect**. As stated above, it cannot be directly measured since only one of the potential outcomes is observed, that is, the university is either gender partial or gender neutral.

In this chapter, we combine the faculty-level models that describe the hiring, initial rank, initial salary, promotion, annual salary adjustment and departure models with a university-level, i.e., an aggregate-level, definition of causal effect. We propose a simulation-based method to generate university-level estimates for a gender neutral world and contrast these to the observed aggregate results in the actual world with some level of gender partiality. The method is based on the fact that the university-level outcomes are functions of individual-level measurements, where we do have replication. Therefore, we can simulate realizations of faculty members' careers over time and aggregate them to get a university-level measure of gender partiality.

Note, the availability of information at the individual level is crucial to the application of this method, since the aggregated causal effect at the university level could not be calculated if the only information available is in aggregated form.

3.1 Challenges in estimating institution-level effects

The methods outlined in section 2.1 provide an array of options to assess gender equity in academic rewards at individual level, i.e., when the unit of analysis is the faculty member. These methods provide options to study outcomes individually or as part of a complex system, the ability to account for the correlation in longitudinal measurements, and, under certain assumptions, the ability to assess causal effects.

However, regardless of the advantages of the methods presented, the analysis of the higher-level causal effects for an aggregated unit such as the institution as a whole presents additional challenges related to the availability of information at the institutional level.

There are two situations in which we might be interested in assessing gender partiality at the institutional level, each with very different purposes.

The first case is multi-institution studies, in which the goal is to estimate the average effect of gender partiality for a geographical region in a period of time. Such studies ask questions like: are gender neutral institutions considered more prestigious; how much more do they invest in their faculty members, when compared to gender partial institutions? This type of study, although not without methodological complications, can be handled with the techniques presented in section 2.1.

The main issue with such studies is that they require knowledge of the gender bias, i.e., whether the institution is gender partial or gender neutral, a priori. If we had such information, we could use the gender partiality status as a binary variable in, for example, a regression model for university-level outcomes on gender partiality and characteristics of the institution. However, classifying institutions as gender-neutral or gender-partial is not trivial. Institutional dynamics are complex and gender partiality might be present at different points of a person's academic career. Real-life gender-neutral institutions might not exist. Using the gender equity perception the institution has of itself might prove to be an inaccurate (perhaps optimistic) portrayal of the actual situation. Trying to assign gender partiality roles to institutions is not only infeasible for a long period of time, but highly unethical.

The second case corresponds to single-institution studies in which we wish to assess the gender equity situation of the institution on its own. In this particular case, the methods presented in section 2.1 are insufficient to assess gender equity. What we would like to do, is to compare the observed institution to itself in a gender neutral scenario. This would allow us to assess whether the observed institution-level outcomes are consistent with gender-neutrality.

The problem with this type of study is that, although we have information for individual faculty members and institution dynamics, we have only one observation for the institution-level outcomes (a sample size of 1), so we only have information for the observed scenario and no information for the institution under the gender neutral scenario, reducing this problem to the estimation of an individual institution-level causal effect, which cannot be directly calculated.

The rest of this chapter will focus on setting up the causal framework for single-institution studies (section 3.2) to later present a method to estimate and conduct inference for such effects (sections 3.3 and 3.4).

3.2 Causal inference framework

Consider the unit of analysis to be an academic institution U composed of $f = 1, ..., F$ faculty members over a period of time $t = 1, ..., T$, and let Z represent the gender partiality status of U, where $Z = 1$ corresponds to an observed (potentially gender-partial) scenario and $Z = 0$ to a gender-neutral scenario.

Based on the individual information on employment status (H_{ft}), salary (S_{ft}) and academic rank (R_{ft}) for faculty member f at time t, we define the following representation and academic rewards outcomes of interest at the university-level $Y = (HD, TR1, TR2, TS)$:

- Hazard of departure of faculty members from the institution, HD, at time t_k:

$$HR(t_k) = \frac{\sum_{f=1}^{F_{t_k}} I(H_{ft_k} = 0)/F_{t_k}}{\sum_{t=t_k}^{T} \left[\sum_{f=1}^{F_{t_k}} I(H_{ft_k} = 0)/F_{t_k} \right]}$$

  where F_{t_k} corresponds to the number of faculty members at the institution at time t_k.

- Total "higher rank" years, i.e, total time faculty members spend at Associate or Full Professor level (as opposed to Assistant level) over the period of time T, as a measure of responsibility, prestige, and success within the institution:

$$TR1 = \sum_{f=1}^{F} \sum_{t=1}^{T_f} I(R_{ft} = \text{“Associate”})$$

$$TR2 = \sum_{f=1}^{F} \sum_{t=1}^{T_f} I(R_{ft} = \text{“Full”})$$

  where the function I takes values 0 or 1 depending on whether the condition $R_{ft} =$ "$rank$" holds, i.e., if faculty member f at time t holds the position "rank".

- Total compensation in dollars awarded to faculty members over the period of time T, as a measure of the investment the institution makes on its faculty:

$$TS = \sum_{f=1}^{F} \sum_{t=1}^{T_f} S_{ft}$$

We are then interested in estimating the "individual" institution's causal effect of gender partiality on the outcomes defined above. This entails constructing a comparison of the potential outcomes $Y(Z = 1)$ and $Y(Z = 0)$, appropriate to the nature of Y. For the outcomes defined above, these causal effects are given by $CE = (HRD, TRL1, TRL2, RTS)$:

- $HRD = \dfrac{HD(Z=1)}{HD(Z=0)}$,

  the hazard ratio of departure from the university comparing the observed to the gender-neutral scenario, assuming proportional hazards over time.

- $TRL1 = \dfrac{TR1(Z=1) - TR1(Z=0)}{\text{Total person years}}$

- $TRL2 = \dfrac{TR2(Z=1) - TR2(Z=0)}{\text{Total person years}}$

  the rates of time in higher rank lost at Full and Associate level, $TRL1$ and $TRL2$ respectively, defined for the common period of observations of the faculty member f under the observed and gender-neutral scenario,

- $RTS = \dfrac{TS(Z=0) - TS(Z=1)}{TS(Z=1)}$

  the relative total compensation, defined for the common period of observations of the faculty member f under the observed and gender-neutral scenario.

3.3 Simulation-based method of estimation

We will follow the next multi-step approach to estimate the institution-level causal effects defined above:

1. Develop a set of nested models for the observed system behavior.
2. Estimate the nested models from the observed data representing the gender partial experience.
3. Develop a model for the gender neutral counterfactual system.
4. Simulate K iterations of the system and its counterfactual.
5. Calculate the average causal effects across iterations.

3.3.1 Develop a set of nested models for the observed system behavior

University U is a complex system in terms of its faculty members careers (see section 1), each one of them defined as a set of transitions from hiring to departure through the hierarchical academic ranks. This system is represented by the following set of nested models (Figure 3.1):

1. the **Arrival Model** represents the joint initial salary (S_0), initial rank (R_0) and hired status (H_0) as a function of gender (G) and the history of characteristics of the faculty member ($\underset{\sim}{X}_0$) (human capital and other variables) up to being hired. Using the standard decomposition of a joint distribution into conditional distributions, this model (for the fth faculty member) can be written as:

$$
\begin{aligned}
[S_{f0}, R_{f0}, H_{f0} = 1 \mid \underset{\sim}{X}_{f0}, G_f] \quad = \quad & [S_{f0} \mid R_{f0}, H_{f0} = 1, \underset{\sim}{X}_{f0}, G_f] \\
& [R_{f0} \mid H_{f0} = 1 \mid \underset{\sim}{X}_{f0}, G_f] \\
& [H_{f0} = 1 \mid \underset{\sim}{X}_{f0}, G_f]
\end{aligned}
$$

Note that information on the last term of the decomposition, corresponding to the probability of being hired given the candidate's characteristics and gender might not be available as a part of administrative records of University U. This information needs to be derived from the hiring process.

2. the **Longitudinal Model** depicts joint salary, rank and hired status over *discrete time*. Note that the model corresponds to the joint history of salary ($\underset{\sim}{S}_T$), rank ($\underset{\sim}{R}_T$) and employment status ($\underset{\sim}{H}_T$) as a function of initial salary, initial rank and initial hired status, as well as the history of characteristics ($\underset{\sim}{X}_T$), productivity ($\underset{\sim}{W}_T$) and gender of the faculty member. This can be written as a transition model in which we assume that the correlation between repeated measurements of the vector of (salary, rank, employment status) exists because past values of the vector influence its current probability distribution.[121] Then, for the fth faculty member:

$$
\begin{aligned}
[\utilde{S}_{fT_f}, \utilde{R}_{fT_f}, \utilde{H}_{fT_f} &= 1 \mid S_{f0}, R_{f0}, H_{f0} = 1, \utilde{X}_{fT_f}, \utilde{W}_{fT_f}, G_f] \\
&= \prod_{t=1}^{T_f} [S_{ft}, R_{ft}, H_{ft} = 1 \mid \utilde{S}_{ft-1}, \utilde{R}_{ft-1}, \utilde{H}_{ft-1} = 1, \utilde{X}_{ft}, \utilde{W}_{ft}, G_f] \\
&= \prod_{t=1}^{T_f} [S_{ft} \mid \utilde{S}_{ft-1}, \utilde{R}_{ft}, \utilde{H}_{ft} = 1, \utilde{X}_{ft}, \utilde{W}_{ft}, G_f] \\
&\quad \cdot [R_{ft} \mid \utilde{S}_{ft}, \utilde{R}_{ft-1}, \utilde{H}_{ft} = 1, \utilde{X}_{ft}, \utilde{W}_{ft}, G_f] \\
&\quad \cdot [H_{ft} = 1 \mid \utilde{S}_{ft-1}, \utilde{R}_{ft-1}, \utilde{W}_{ft}, \utilde{W}_{ft}, G_f]
\end{aligned}
$$

Conditioning on the history of the faculty members' characteristics, productivity and gender, this model further specifies that:

- current salary depends on salary history up to the previous time point, as well as current rank and employment status,
- current rank depends not only on rank history up to the previous time point and employment status, but it might also depend on salary history,
- employment status, that is, the decision or imposition to leave the institution, depends on the history of salary and rank up to the previous time point given the person was still employed then.

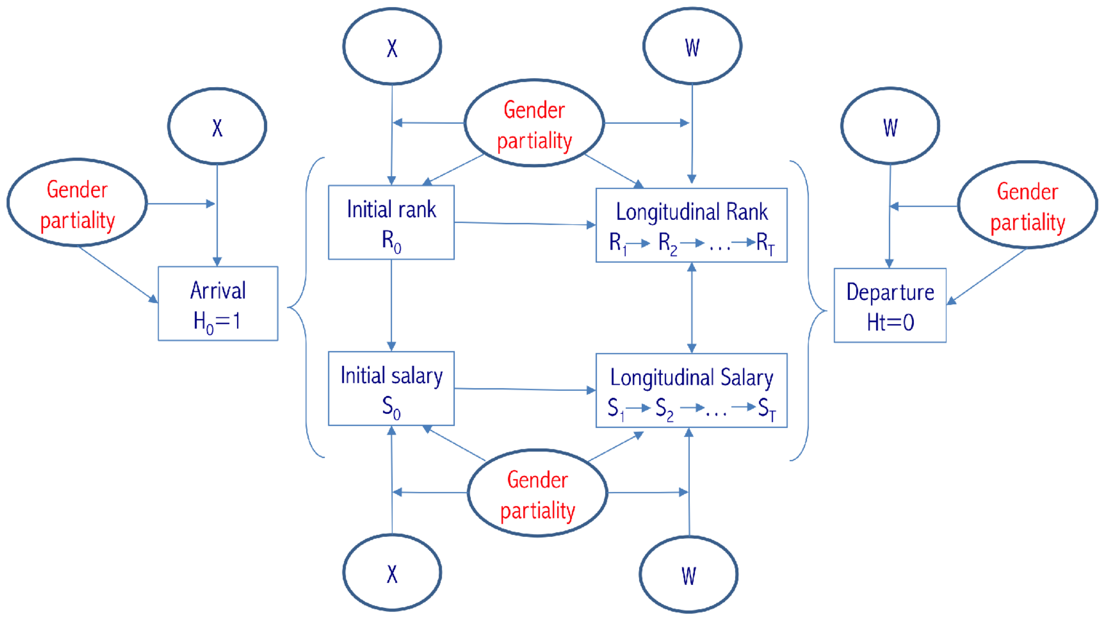


Figure 3.1: Observed system behavior for University U. Faculty members are hired and assigned ranks based on qualifications. Initial salary, as well as longitudinal rank and salary depend on history of both rank and salary, as does departure from the institution. Each node is conditioned on still being employed. Any of these nodes can be affected by gender partiality.

3.3.2 Estimate nested models from the observed data

Given the system structure described above, we implement the approach with specification of the forms of the conditional models and use standard estimation methods as follows:

1. $[H_0 = 1 \mid \underset{\sim}{X}_0, G]$, hiring model, fit on applicant pool information[3].

2. $[R_0 \mid H_0 = 1, \underset{\sim}{X}_0, G]$, initial rank model, fit on newly hired faculty data.

3. $[log(S_0) \mid R_0, H_0 = 1, \underset{\sim}{X}_0, G]$: log initial salary model, fit on newly hired faculty.

[3]Note that applicant pool might not be available in which case this step is omitted

4. $[P_t = 1 \mid \underset{\sim}{S}_{t-1}, \underset{\sim}{R}_{t-1}, \underset{\sim}{H}_t = 1, \underset{\sim}{X}_t, \underset{\sim}{W}_t, G]$: rank-specific models for promotion to the following rank, fit on the set of faculty members eligible to be promoted.

5. $\left[\log\left(\frac{S_t}{S_{t-1}}\right) \middle| \underset{\sim}{S}_{t-1}, \underset{\sim}{R}_t, \underset{\sim}{H}_t = 1, \underset{\sim}{X}_t, \underset{\sim}{W}_t, G\right]$: model for log change in salary with respect to the previous year, fit to the entire dataset excluding newly hired faculty members.

6. $[H_t = 0 \mid \underset{\sim}{S}_t, \underset{\sim}{R}_t, \underset{\sim}{H}_{t-1} = 1, \underset{\sim}{X}_t, \underset{\sim}{W}_t, G]$: departure model, fit to the entire dataset.

   Different models can be estimated for departure at the end of first year and departures any time after that, since these departure processes might depend on different covariates.

Models should include all relevant covariates and all gender main effects and interactions thought to potentially operate at the institution. Salary models should include random intercepts for any clustering structures present at the university (departments, for example) and take into account correlation over time.

The type and specific functional forms will depend on the particular mechanisms in place, available data, and model diagnostics. Reasonable models include logistic regression for hiring, adjacent category logit models for initial rank, and log-link logistic regressions for promotions and departures.

This step will produce a vector of estimated coefficients $\hat{\theta}_1$ for all relevant variables included in the career models for the observed, potentially gender-partial scenario.

3.3.3 Develop a model for the counterfactual system behavior

In University U's counterfactual system behavior as a gender-neutral institution, there would be no influence of gender partiality at any of the steps of the system (Figure 3.2). This manifests necessarily as a system in which there are no differences in the representation and academic rewards outcomes for female and male faculty.

We need to specify the effect of each of the covariates on the institution-level outcomes in the gender-neutral counterfactual, i.e., the coefficients of the covariates in the nested models that would be in place if university U were gender neutral[4].

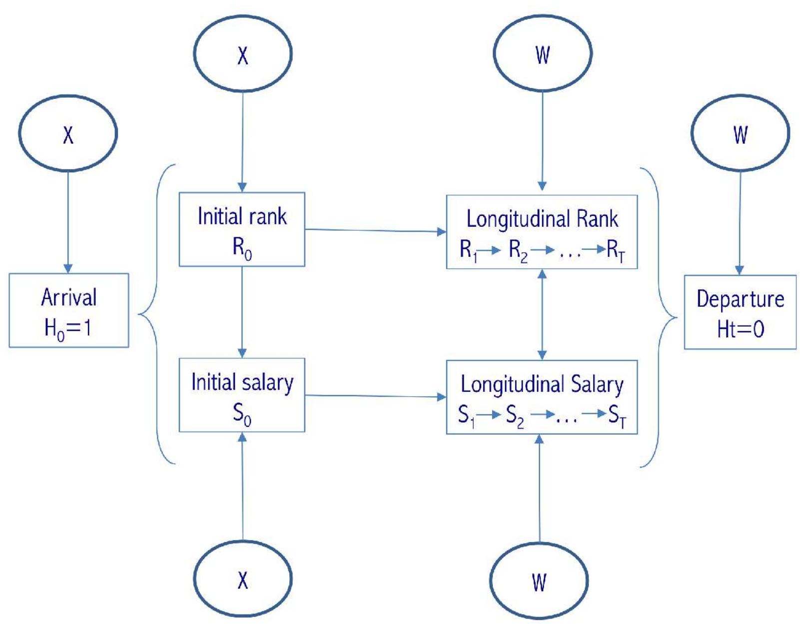


Figure 3.2: Gender-neutral counterfactual system behavior for University X. There is no effect of gender partiality at any point of a faculty member's academic career.

[4]This has been called "return on the covariates" in the gender equity literature

This step then produces a vector of estimated coefficients $\hat{\theta}_0$, for all relevant variables included in the career models, under the chosen gender-neutral scenario.

The choice of the appropriate gender-neutral form depends on the particulars of each institution. The choices are the same ones confronted by the decomposition techniques presented in section 2.1.4. Then, $\hat{\theta}_0$ can take any of the following structures:

1. Female: set $\hat{\theta}_0$ to the coefficients of the career models fit for female faculty only. This assumes male faculty are overpaid in the observed scenario.

2. Male: set $\hat{\theta}_0$ to the coefficients of the career models fit for male faculty only. This assumes female faculty are underpaid in the observed scenario.

3. Pooled: set $\hat{\theta}_0$ to the coefficients of the career models fit for all faculty members, excluding gender from the models. This assumes male faculty are overpaid and female faculty are underpaid at the same time in the observed scenario, and the coefficients for the gender- neutral scenario are derived by ignoring gender.

4. Weighted: set $\hat{\theta}_0$ to the coefficients of the career models fit for all faculty members, with the gender coefficients changed to zero. This assumes male faculty are overpaid and female faculty are underpaid at the same time in the observed scenario, and the coefficients for the gender-neutral scenario are adjusted for gender.

3.3.4 Simulate K realizations of the system and its counterfactual

Once both system and counterfactual parameters have been estimated ($\hat{\theta}_1$,$\hat{\theta}_0$), we will use this information to generate K realizations of each faculty member's career under the observed and gender-neutral scenario ($Career_{fk}(\hat{\theta}_1)$,$Career_{fk}(\hat{\theta}_0)$), following the models outlined in sections 3.3.1 and 3.3.3.

Depending on the behavior chosen for the counterfactual, gender neutrality will have an effect on both male and female faculty members (female, male or pooled structure), or female faculty members only (weighted structure). We can then choose to run the simulation on all faculty members or female faculty members only. In cases where both females and males are included, results will be given separately for each.

A single iteration k of this simulation will follow the steps shown in Figure 3.3. Each step is run twice with parameters $\hat{\theta}_1$ and $\hat{\theta}_0$ separately, to generate realizations of the observed and gender-neutral scenario:

1. If information for the pool of applicants to an academic position is available, generate hiring decisions as random Bernoulli draws with probability p_H estimated using the hiring model.

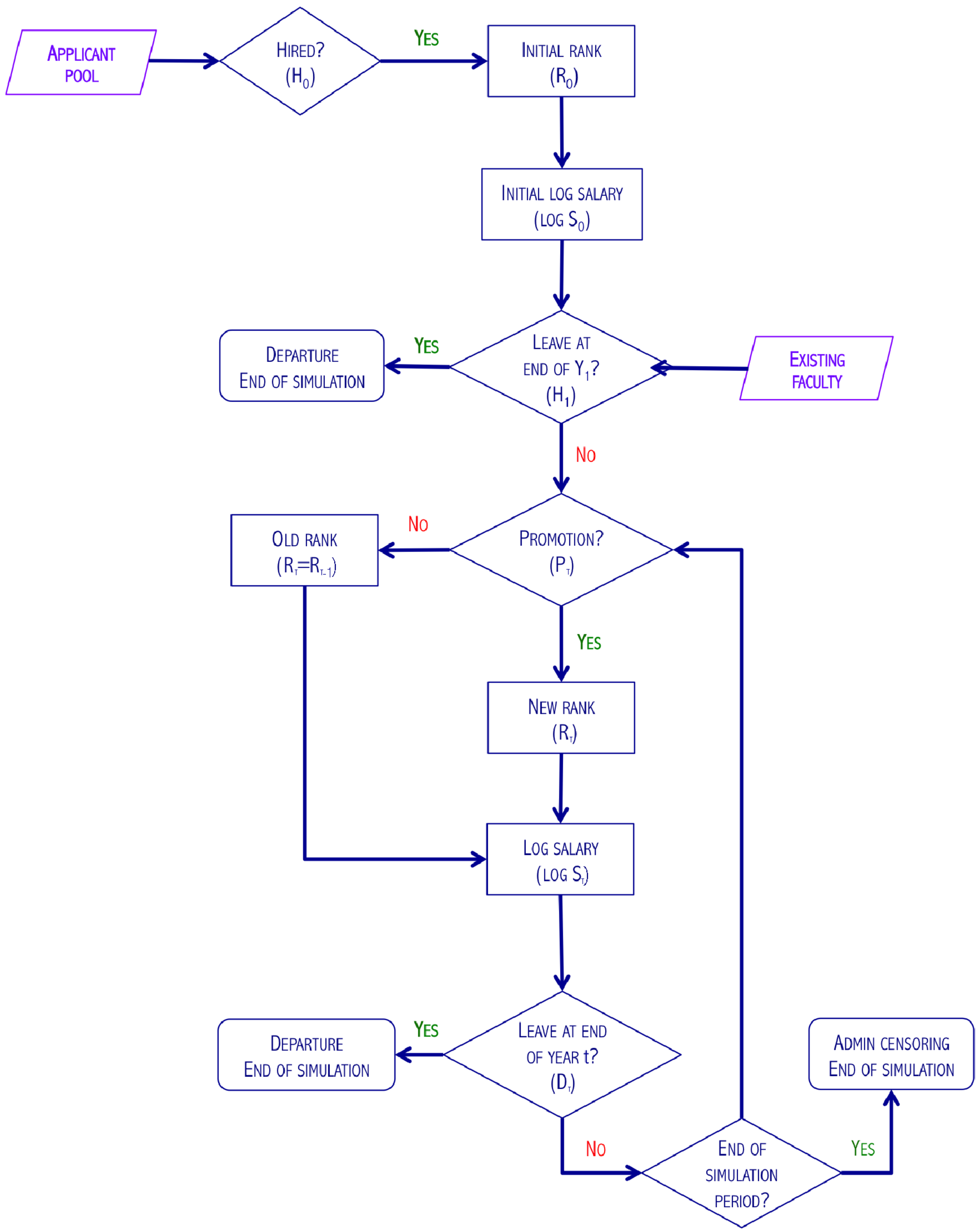


Figure 3.3: Academic career simulation procedure. Given a pool of applicants and a set of existing faculty members, the procedure simulates the careers of faculty from hiring to departure or end of simulation, by recreating the hiring, promotion and salary assignment mechanisms based on the fitted career models.

For **newly hired faculty**:

2. Generate initial rank:

   If positions at the university are advertised as open rank, and rank is assigned to the new faculty member as a function of human capital variables, generate initial rank as a random draw from the ranks Assistant, Associate or Full professor, with probabilities (p_{R1},p_{R2},p_{R3}) estimated using the initial rank model.

   If the university hires faculty members for specific ranks, use observed initial ranks.

3. Generate initial salaries:

   Initial log salaries are generated from the initial log salary model, and need to be transformed back to their original scale. Assuming the log-transformed initial salaries are Gaussian with variance σ^2, we could estimate the mean initial salary as $exp(\widehat{log(S_0)} + \hat{\sigma}^2/2)$. However, the lognormality and constant variance assumptions are very strong and rarely hold. Instead, we propose a method based on Duan's smearing estimator to produce a non-parametric estimation of the initial salaries.[122]

   Assuming a mixed effects model with a random intercept at department level and different residual variances by department, then:

   (a) Get the predicted mean log initial salary from the initial salary model, including the department random effects:

$$\widehat{log(S_0)} + \hat{d}_i \quad \text{where } d \sim N(0, \tau^2),\ i = 1, ..., D$$

where D represents the number of departments.

(b) Add noise at faculty member level $\hat{\epsilon}_{if}$.

$\hat{\epsilon}_{if}$ is selected at random from the distribution of standardized female residuals (for the observed scenario) or the distribution of standardized male residuals (for the gender-neutral scenario), on the log scale, from the initial salary model, and unstandardized to the department of faculty member f.

$$\widehat{log(S_0)} + \hat{d}_i + \hat{\epsilon}_{if}$$

(c) Generate initial salary as

$$\widehat{S_{f0}} = e^{\widehat{log(S_0)} + \hat{d}_i + \hat{\epsilon}_{if}} \cdot \frac{1}{F_{i0}} \sum_{f=1}^{F_{i0}} e^{\hat{\epsilon}_{if}}$$

where F_{i0} is the number of newly hired female (observed scenario) or male (gender-neutral scenario) faculty members at department i.

For both **new and existing** faculty:

4. Generate initial departure decisions:

If there are enough departures at the end of the first year in the dataset so that a separate departure model was estimated for the initial departure, or if there is sufficient information in the dataset to predict initial departures using a general departure model, generate initial departures as a random draw from a Bernoulli distribution with probability p_D, estimated from the departure model.

If there are not enough departures at the end of the first year, assume no faculty member leaves the institution at this time.

Then, **conditional on staying at the institution**:

5. Generate promotion decisions as random Bernoulli draws with probabilities (p_{Assoc} ,p_{Full}), estimated from the rank-specific promotion models.

6. If the faculty member is promoted, assign next academic rank, otherwise he/she remains in current rank.

7. Generate new salary following the smearing-based procedure outlined above:

   (a) Get the predicted mean log change in salary from the change in salary (*saldt*) model, including any random effects if present.

   Assuming a mixed effects model with time nested in faculty members, and faculty members nested in departments:

$$\widehat{saldt}_{ft} + \hat{d}_i + \hat{d}_{if}$$

   where $d_i \sim N(0, \tau_1^2)$ and $d_{if} \sim N(0, \tau_2)^2$ correspond to the department and person random effects respectively.

(b) Add noise at time/faculty member level $\hat{\epsilon}_{ift}$

$\hat{\epsilon}_{ift}$ is selected at random from the distribution of standardized female residuals (for the observed scenario) or the distribution of standardized male residuals (for the gender-neutral scenario),on the log scale, from the change in salary model, and unstandardized to the department of faculty member f.

$$\widehat{saldt}_{ft} + \hat{d}_i + \hat{d}_{if} + \hat{\epsilon}_{ift}$$

(c) Generate new log salary as:

$$\widehat{log(S_{ft})} = \widehat{saldt}_{ft} + \hat{d}_i + \hat{d}_{if} + \hat{\epsilon}_{ift} + \widehat{log(S_{f,t-1})}$$

(d) Generate new salary as:

$$\widehat{S_{ft}} = e^{\widehat{log(S_{ft})}} \frac{1}{F_i} \textstyle\sum_{f=1}^{F_i} e^{\hat{\epsilon}_{jft}}$$

where F_i is the number of female (observed scenario) or male (gender-neutral scenario) faculty members at department i.

8. Generate departure decisions as random Bernoulli draws with probability p_{if}, estimated from the departure model.

9. Repeat steps 5-8 a total of $T - 1$ times, for a simulation time period of length T.

3.3.5 Estimate average causal effects

After the simulation above is done, we will have K realizations of the institution under the observed and gender-neutral scenarios. We can then use the individual faculty member's information to calculate institution-level causal effects for each realization $k = 1, ..., K$:

- $HRD_k = \dfrac{HD_k(Z=1)}{HD_k(Z=0)}$

- $TRL1_k = \dfrac{TR1_k(Z=1) - TR1_k(Z=0)}{\text{Total person years}}$

- $TRL2_k = \dfrac{TR2_k(Z=1) - TR2_k(Z=0)}{\text{Total person years}}$

  where $TRL1_k$ and $TRL2_k$ are defined for the common period of observations of the faculty member f under the observed and gender-neutral scenario,

- $RTS_k = \dfrac{TS_k(Z=0) - TS_k(Z=1)}{TS_k(Z=1)}$,

  defined for the common period of observations of the faculty member f under the observed and gender-neutral scenario.

Finally, the average causal effects ($\overline{CE}$) are given by:

$$\overline{CE} = (\overline{HRD}, \overline{TRL1}, \overline{TRL2}, \overline{RTS})$$

Where:

$$\overline{HRD} = \frac{1}{K}\sum_{k=1}^{K} HRD_k$$

$$\overline{TRL1} = \frac{1}{K}\sum_{k=1}^{K} TRL1_k \qquad \overline{TRL2} = \frac{1}{K}\sum_{k=1}^{K} TRL2_k$$

$$\overline{RTS} = \frac{1}{K}\sum_{k=1}^{K} RTS_k$$

3.4 Inference for estimated higher-order causal effects

We have defined and demonstrated the estimation of the point estimates of the causal effect of gender-partiality on the institution. This section addresses variance and confidence interval estimation.

Since the variance of the average causal effects has no closed form solution, we will use resampling procedures to calculate it. We use the case bootstrap procedure where individual faculty members are resampled conditional on gender and department structure, in order to maintain the distribution of these characteristics in each bootstrap sample.

We will then run the simulation-based method to estimate the average causal effects on each of a total of B bootstrap samples (e.g. 250 samples). This implies refitting the nested career models, simulating the system and its counterfactual K times and estimating $\overline{HRD}$, $\overline{TRL1}$, $\overline{TRL2}$ and $\overline{RTS}$ as averages across the K realizations of the institution, for each bootstrap sample.

This process is computationally intensive, since it entails refitting the nested career models and simulating academic careers a total of $2KB$ times: 2 scenarios (observed and gender-neutral), K realizations, B bootstrap samples. We introduce then the following statistical and computational approaches to reduce computational time:

1. Use of one-step estimators when refitting the career models during bootstrap.

   In this approach, proposed by Moulton and Zeger (1991), we supply the algorithm with the model coefficients estimated for the original dataset as initial values and take one Newton-Raphson iteration in such direction, instead of refitting the career models to convergence.[123]

   This method not only reduces computation time, but also provides more stable bootstrap estimators in the case of extreme bootstrap samples.

2. Model the variance of the estimated effect as a function of the inverse number of realizations.

   In theory, we would like to simulate an infinite number of realizations ($K \to \infty$) of University U under the observed and gender-neutral scenarios, which is, of course, not feasible. Instead, we can run a reduced number of realizations

$K_r < K$ and model the variance of the estimated effect as:

$$E[Var(\overline{CE_i})_k] = \beta_0 + \beta_1 \frac{1}{k}$$

Where $\overline{CE_i}$ represents one of the 4 effects of interest HRD, $TRL1$, $TRL2$ or RTS, $k = 2, ..., K_r < K$ and $Var(\overline{CE_i})_k$ is calculated over k realizations of U. In this model, β_0 can be interpreted as the expected variance of the effect when $1/k$ is zero, that is, when $k \rightarrow \infty$, so we can use $\hat{\beta}_0$ as an estimator for the desired variance.

3. From the computational point of view, the use of cloud computing and parallel programming reduces running time greatly. To further reduce computation time, the simulation method algorithm combines R and C++ code, through Rcpp, which resulted in a 75% reduction in running time.

Once variances are estimated, we can estimate confidence intervals for the causal effects that will let us assess whether the observed institution is significantly different from its gender neutral counterpart. Intervals that contain 1 (in the case of the average HRD) and intervals that contain zero (for average TRL and RTS) indicate no statistically significant departures from gender-neutrality.

Code to implement the simulation-based method and perform the bootstrap variance calculation, along with a toy dataset example, will be released in the future as an R package.

Chapter 4

Assessing sources of information in complex designs

Once institution-level effects have been estimated, and if any departures from gender neutrality exist, the next step in the analysis is to identify the determinants in the academic career that most contribute to gender partiality. That is, how do gender disparities observed in the nested career models affect the estimated institution-level causal effect? In this chapter we present a regression-based method to identify the sources of variability of estimated effects in complex designs, as a function of existing differences within their individual components.

The desire to partition sources of variation within the complex system is one that arises in several different applications. The methods presented here are motivated by the gender equity problem but also apply to other complex systems including complex study designs, e.g. a stepped wedge design. In a stepped wedge, a researcher may be interested in the separation of cross-sectional and longitudinal evidence about a treatment effect. Section 4.1 presents a small digression from the overall theme of this dissertation to introduce the stepped wedge design and set up the method to assess the contributions to estimates of a complex intervention effect. Sections 4.1.2 and 4.1.3 present the proposed method and simulation results within the stepped wedge paradigm. We finish this chapter by extending the proposed method to more complex systems, and discuss its application in the context of studying gender partiality in academia 4.2.

4.1 Motivating application : Stepped wedge designs

Designs where intervention assignment is done at a group level are known as cluster designs. An example of such a design is the assignment of an intervention to a whole village, as opposed to assigning the intervention to the individual villagers. This type of design is frequently used when individual assignment is not feasible, there is the potential for interference or contamination between individuals, or when we are interested in assessing a higher-level effect, such as at the village level.

There are several ways in which assignment of intervention can be done at the cluster level. The stepped wedge design is a special case of the crossover design where all the clusters receive control followed by the intervention. However, the duration of time the clusters are observed during the control and intervention period differs and this defines the stepped wedge design.[124] Specifically, at the beginning of the study all clusters are assigned a time when they will cross-over to the intervention, such that by the end of the trial, all the clusters are receiving the intervention.[125] Table 4.1 shows an example of a complete stepped design[1].

	Time				
Cluster	1	2	3	4	5
1	0	1	1	1	1
2	0	0	1	1	1
3	0	0	0	1	1
4	0	0	0	0	1

Table 4.1: Example of complete stepped wedge design. At time 1, all clusters start out in the control arm of the study (0). At subsequent times, clusters start receiving the intervention (1) until all clusters are exposed (time 5). Measurements are conducted at each of the 5 time points.

The stepped wedge design is useful in cases where the intervention is difficult to implement within a large number of clusters simultaneously. However, since more clusters receive the intervention towards the end of the study, the intervention effect "might be confounded with underlying temporal trends".[124] It is then important to assess the effect that cross-sectional and longitudinal information have on the overall intervention effect.

[1]Note that there are other possibilities, such as incomplete designs, as discussed in Hemming et al. (2014). We shall focus here on the complete design.

4.1.1 Model and contrasts

In cluster designs, it is common that the individual measurements are not independent. In order to take this within-cluster correlation into account, we set up the following mixed model for a Gaussian response y_{ijt} of interest measured on individuals within clusters:

$$y_{ijt} = \beta_0 + b_{0i} + b_{0ij} + (\beta_1 + b_{1i})Z_{it} + \beta_S X_S + \epsilon_{ijt} \tag{4.1}$$

Where:

- i denotes cluster, $i = 1, ..., N$
- j denotes person, $j = 1, ..., J$
- t denotes time, $t = 1, ..., K$
- Z_{it} is an intervention indicator: it takes the value 0 if cluster i at time t is in the control arm, and it takes the value 1 if cluster i at time t is receiving the intervention
- X_S and β_S are a matrix and vector respectively of variables and coefficients for a natural spline of S degrees of freedom on time
- b_{0i} and b_{0ij} correspond to random intercepts at cluster and person level, with $b_{0i} \sim N(0, \sigma_c^2)$ and $b_{0ij} \sim N(0, \sigma_p^2)$
- b_{1i} corresponds a random slope on the treatment indicator, with $b_{1i} \sim N(0, \sigma_I^2)$
- ϵ_{ijt} are the residual errors, $\epsilon_{ijt} \sim N(0, \sigma_e^2)$

Additionally:

- For a particular cluster i, K_{0i} and K_{1i} denote the number of observations before and after cluster i was assigned to the intervention, respectively.

- For a particular time t, N_{0t} and N_{1t} denote the number of clusters assigned to control and intervention, respectively.

We are interested in investigating the contribution of cross-sectional and longitudinal information to the overall estimated intervention effect. That is, we seek to assess the contributions of two sources of evidence to the final effect estimate: (1) the differences in mean response between intervention and control at each time point; and (2) the mean differences after versus before the crossover for each cluster. We call these the "cross-sectional" and "longitudinal" contrasts, respectively and define them as follows.

4.1.1.1 Cross-sectional contrasts

For a particular time t, let $\bar{y}_{C0t}$ and $\bar{y}_{C1t}$ be the mean response (at person level) for clusters assigned to control and intervention, respectively. Then:

$$\bar{y}_{C1t} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{J} y_{ijt}Z_{it}}{J \cdot N_{1t}} \qquad \bar{y}_{C0t} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{J} y_{ijt}(1 - Z_{it})}{J \cdot N_{0t}}$$

The cross-sectional contrasts are given by $\bar{y}_{C1t} - \bar{y}_{C0t}$ with variance:

$$Var(\bar{y}_{C1t} - \bar{y}_{C0t}) = \frac{\sigma_I^2}{N_{1t}} + \left(\frac{N}{N_{0t} \cdot N_{1t}}\right) \cdot \left(\sigma_c^2 + \frac{\sigma_p^2}{J} + \frac{\sigma_e^2}{J}\right)$$

4.1.1.2 Longitudinal contrasts

For a particular cluster i, let $\bar{y}_{L0i}$ and $\bar{y}_{L1i}$ be the mean response before and after the cluster is assigned to the intervention. Then:

$$\bar{y}_{L0i} = \frac{\sum_{t=1}^{K} \sum_{j=1}^{J} y_{ijt}(1 - Z_{it})}{J \cdot K_{0i}} \qquad \bar{y}_{L1i} = \frac{\sum_{t=1}^{K} \sum_{j=1}^{J} y_{ijt} Z_{it}}{J \cdot K_{1i}}$$

The longitudinal contrast for cluster i is given by $\bar{y}_{L1i} - \bar{y}_{L0i}$, with variance:

$$Var(\bar{y}_{L1i} - \bar{y}_{L0i}) = \sigma_I^2 + \frac{K}{J \cdot K_{0i} \cdot K_{1i}} \sigma_e^2$$

The behavior of these variances is important to set up simulation parameters that would allow us to assess which components, cross-sectional or longitudinal, contribute more to the estimation of the treatment effect. We created a shiny app that allows a researcher to assess the effects of σ_I^2, σ_c^2, σ_p^2, σ_e^2, K, K_{0i}, K_{1i}, N, N_{0t}, N_{1t} and J on the magnitude and behavior of the variance of the contrasts. The shiny app can be accessed at `https://francisabreu.shinyapps.io/Variance_of_Contrasts/`

Appendix 8.1 provides further comments on the behavior of these variances.

4.1.2 Method

We are interested in assessing the contribution of the cross-sectional and longitudinal contrasts on the treatment effect β_1. We can do this by conditioning the estimated model coefficients on the cross-sectional and longitudinal contrasts and analyzing the reduction in the variance of β_1 and the weight of each individual contrast on the value of β_1.

Let us consider the general form of a linear mixed model in matrix form as

$$Y = X\beta + Z\gamma + \epsilon$$

where Y is the vector of responses, β is the vector of unknown parameters associated with the fixed effects, X and Z are known matrices and γ and ϵ are the unobservable random effects and residuals respectively. We can estimate β as:

$$\hat{\beta} = (X^t V^{-1} X) X^t V^{-1} Y = MY$$

where V is the variance-covariance matrix of Y.

Let us consider AY a vector of contrasts. If AY corresponds to cross-sectional contrasts, it will contain, for each time t, the difference in mean response between the clusters assigned to the intervention and the clusters assigned to control. If AY corresponds to longitudinal contrasts, it will contain, for each cluster i, the difference in mean response between the period of time when cluster i was assigned to the intervention and the period of time when it was assigned to control.

Assuming Y follows a multivariate Gaussian distribution $Y \sim MVN(X\beta, V)$, the variance of $\hat{\beta}$ is given by $Var(\hat{\beta}) = MVM^t$

The conditional distribution of $\hat{\beta}$ on AY is then multivariate normal

$$\hat{\beta} \mid AY \sim MVN(E[\hat{\beta} \mid AY], Var[\hat{\beta} \mid AY])$$

with:

$$E[\hat{\beta} \mid AY] = \beta + MVA^t(AVA^t)^{-1}(AY - AX\beta) \tag{4.2}$$

$$Var[\hat{\beta} \mid AY] = MVM^t - MVA^t(AVA^t)^{-1}AVM \tag{4.3}$$

We can then use the coefficients of the contrasts AY in equation 4.2, that is, $\lambda = MVA^t(AVA^t)^{-1}$ to assess the effect of each individual contrast on the treatment effect, and compare the variances obtained using equation 4.3 to the variance of the treatment effect $Var(\hat{\beta})$ to assess the percent reduction in variance attributable to cross-sectional and longitudinal components.

The following section presents a simulation study in which we apply the method described.

4.1.3 Simulation study

We generated data for a Gaussian response as measured in a stepped wedge design following model 4.1, for combinations of several values of the parameters J, K, N, β_1, σ_C^2, σ_I^2. Model intercept was assumed to $\beta_0 = -2.5$, $\sigma_e^2 = 1$, the person-level variance $\sigma_P^2 = \sigma_c^2 \cdot N$ and, if a random slope on treatment is present, $\sigma_I^2 = (\beta_1/2)^2$. Appendix

8.2 lists all values considered for each parameter.

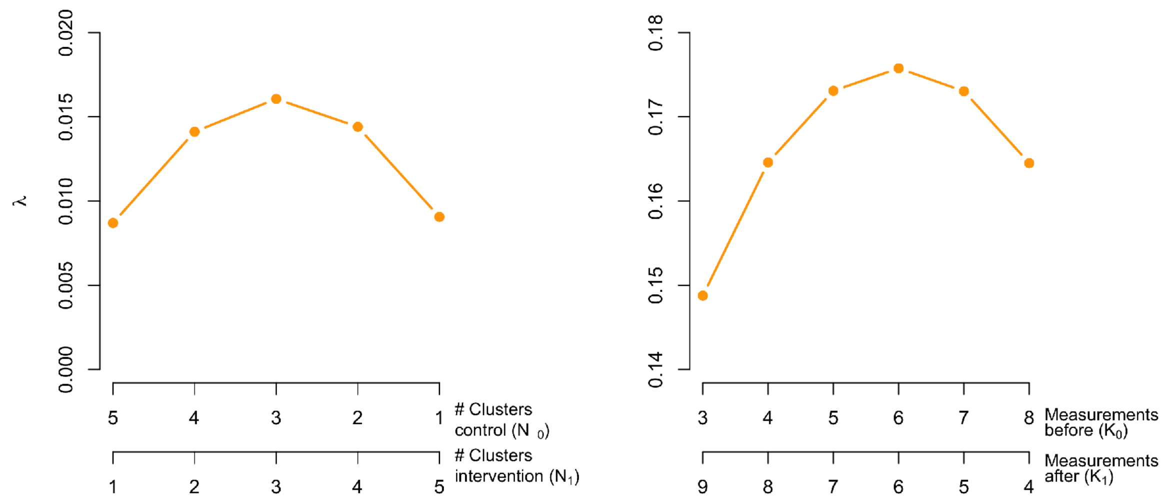
Results below correspond to a scenario with $N = 6$ clusters, with 10 individuals per cluster, measured over 12 months. There is a linear temporal trend in the responses with $\beta_1 = -0.5$ and $\sigma_C^2 = 0.5$, with a random slope on the treatment effect. Results for other scenarios will be made available as a shiny app in the future.

Table 4.2 shows the variance of $\hat{\beta}_1$, overall and conditioning on cross-sectional and longitudinal contrasts for models with 5 different degrees of freedom for the natural spline on time. We can observe that the variance of $\hat{\beta}_1$ is greatly reduced ($\sim$ 99.9%) when conditioning on the longitudinal contrasts implying that, in the design considered, the overwhelming fraction of information comes from the longitudinal component of the design.

		Conditioning contrast	
S	Overall	Cross-sectional	Longitudinal
0	0.0107	0.0105	0.000007
1	0.0242	0.0228	0.000007
2	0.0246	0.0232	0.000007
4	0.0272	0.0255	0.000007
8	0.0274	0.0257	0.000007

Table 4.2: Variance of $\hat{\beta}_1$: overall and conditioning on cross-sectional and longitudinal contrasts. S corresponds to the degrees of freedom of the natural spline in model 4.1

Figure 4.1 shows the λ coefficients for cross-sectional and longitudinal contrasts, for the model with $S = 4$. Models with a different number of degrees of freedom show similar patterns.



(a) Cross-sectional **(b)** Longitudinal

Figure 4.1: Coefficients λ for cross-sectional and longitudinal contrasts

The λ coefficients for cross-sectional contrasts have higher values for months closer to the middle of the study. This is expected since this corresponds to periods of time with a more balanced number of clusters in control/intervention and are expected to have a bigger influence on the estimated treatment effect. The λ coefficients for the longitudinal contrasts show that clusters which receive the intervention closer to the middle of the study have larger coefficients. This means that large differences in mean response for clusters with more balanced number of observations in the control /intervention period have a bigger influence on the estimated treatment effects, as expected.

4.2 Assessing sources of information in complex systems

The method presented in section 4.1.2 can be extended to systems more complex than the stepped wedge design by using regression models to assess the association of the overall estimated higher-level causal effect with differences found at each of its components. In the example of gender equity in academia, this means assessing the relationship of the estimated institution-level causal effects to gender disparities found at hiring, initial rank, initial salary, promotions, annual salary adjustments and departures.

In the Gaussian case and in examples like the stepped wedge design presented in sections 4.1.2 and 4.1.3, there are closed forms for the variances and regression coefficients. For the complex system of academia however, the causal effects of interest are complex, non-linear functions of parameters for the observed and the gender neutral scenario $\theta = (\theta_1, \theta_0)$.

Let CE be the causal effects of interest, then we can express the estimated causal effects as a complex function of history of salary, rank, and employment under the observed and gender neutral institution as:

$$\widehat{CE}[(\underset{\sim}{S}, \underset{\sim}{R}, \underset{\sim}{H})(\theta_1), (\underset{\sim}{S}, \underset{\sim}{R}, \underset{\sim}{H})(\theta_0)]$$

The method to assess the sources of information in this case is made easier by using the power of bootstrapping. For a number B of bootstrap samples (e.g. 250), we obtain the estimated causal effects $\overline{CE}^{(b)}$ and the save the **gender coefficients** of the career models $\theta_{1Gj}^{(b)}$ for the observed scenario[2], with $j = 1, ..., J$ representing one of the career steps (hiring, initial rank, initial salary, promotion to associate, promotion to full professor, annual salary adjustments, and departures).

We can then regress $\overline{CE}^{(b)}$ on such gender coefficients and their pairwise interactions as:

$$\overline{CE}^{(b)} \approx \sum_{j=1}^{J} \gamma_j \theta_{1Gj}^{(b)} + \sum_{j,j'}^{J} \gamma_{jj'} \hat{\theta}_{1Gj}^{(b)} \hat{\theta}_{1Gj'}^{(b)}$$

The relative sizes of the γ coefficients and their t-statistics will then inform us about the components that influence the estimated causal effects the most. We present results for this method applied to the gender equity example in section 5.4 for simulated data and in Chapter 6 for University X.

[2]Remember that θ_1 is the vector of coefficients of all variables included in the career model for the observed scenario. The gender coefficients are then a subset of θ_1, which we call θ_{1G}

Chapter 5

Simulation study

This chapter presents simulations that assess the performance of the methods outlined in chapters 3 and 4, under different gender partiality scenarios.

The simulation is conducted in two steps. First, we generate faculty careers within a hypothetical university with known mechanisms of gender partiality. For each repetition of the university, we estimate the vector of effects ($\overline{CE}$) and study their distributions, as well as the sources of information that contribute to the higher-level causal estimators across repetitions of the simulation.

5.1 Data generation under known mechanisms of bias

For a defined time period, we generate a pool of existing faculty members and a pool of applicants, then simulate the hiring process and academic career of the faculty members over time. This can be done by formulating a hypothesis on the dynamics of the institution, with the following elements.

- Size and characteristics of the institution at the start of the simulation period:

  Gender, full time equivalent salary, change in salary with respect to previous year, rank, and time in current rank need to be included as characteristics in the dataset. Optional characteristics may include department, year of hiring, age, degree, race, and productivity measures, among others.

- Size and characteristics of the applicant pool for each year of the simulation:

  Characteristics to be included should be those thought to determine hiring decisions, for example, desirability measures, department of application, gender, degree, age, race, among others. These characteristics may reflect market conditions or recruitment strategies of the institution.

- Hiring model: Model for the probability of being hired given applicant's characteristics. This model should be sensible to the size of the institution. The model can be data driven or coefficients can be set according to university policy or other considerations.

- Career models: Models for initial rank, initial salary, promotions, annual salary adjustments and departure from the institution, as defined in section 3.3.2. The gender coefficients of these models is set to reflect a specified level of bias. These models can be data driven or coefficients can be set according to university policy or other considerations.

5.1.1 Simulation parameters

We generate 500 datasets under four different gender partiality scenarios: neutral (GN), low gender partial (LGP), medium gender partial (MGP) and high gender partial (HGP), further described in page 82. The following parameter values were used to generate the data:

- Time period: 2005-2014 (10 years).
- Institution size: 1000 faculty members at 2005.
- Existing faculty members:

  Distribution of faculty members by gender, race, rank, department and degree reflect US medical school national faculty distributions[126][1]. Faculty are classi-

[1] Table 14table18.xlsx for degree, excluding Dentistry, Other Health Professions, Social Sciences, Veterinary Sciences, All Others, and Public Health.

fied as recently promoted or not by a random draw from a Bernoulli distribution, with probabilities of promotion of $p_F = 0.095$ for female faculty and $p_M = 0.119$ for male faculty.

Current age is assumed to have different means by gender, race, department category (basic or clinical) and rank, following Alexander and Liu.[127] Individual ages are a random draw from a right skew normal distribution with variance=81 and skewness=10, constrained to be between 18 and 90.

Age-at-hire is assumed right skew normal with mean 37.8 years following Alexander and Liu,[127] variance=42.25 and skewness=10, constrained to be between 18 and 90. Age-at-hire is also constrained to be smaller than age, and for professors not newly promoted it is further constrained to be smaller than age-1.[2]

Time in current rank for assistant professors is generated given age and age-at-hire. For newly promoted faculty, it is obtained as a random draw from an uniform distribution $U(\frac{1}{12}, \frac{11}{12})$ while for assistant and full professors not promoted, times are drawn from exponential distributions, respectively $exp(1/6)$ and $exp(1/9)$, and constrained to be larger than 1 and smaller than the time the faculty member has been at the institution.

Salaries ($ftesal$) are assumed to be lognormal with variance for the logged salaries of 0.004 and different means by degree, rank and department, following

[2] Since the person was not promoted, their time in rank has to be greater than 1

the model:

$$
\begin{aligned}
E[log(ftesal)] \quad = \quad & 12.5 - 0.05 \cdot female + 0.25 \cdot MD + 0.25 \cdot PhD \\
& -0.25 \cdot Associate - 0.5 \cdot Assistant \\
& -0.4 \cdot BasicSciences - 0.4 \cdot FamilyMedicine \\
& -0.4 \cdot InternalMedicine - 0.5 \cdot Neurology - 0.35 \cdot ObGyn \\
& -0.35 \cdot Other - 0.45 \cdot Pathology - 0.50 \cdot Pediatrics \\
& -0.50 \cdot Psychiatry - 0.2 \cdot Radiology - 0.2 \cdot Surgery
\end{aligned}
$$

- Applicant pool:

  The size of the applicant pool by year is a random draw from a Poisson distribution with mean 5% of the number of medical school graduates.[128]

  The department of application reflects the distribution of active residents by GME specialty for 2013-2014.[129] Gender and race reflect the national distribution of medical graduates[130,131] over time, and race and is assumed to be unrelated to the rest of the covariates. Degrees are assumed to depend on gender and department of application, reflecting existing faculty distribution. Age is assumed to be right skew normal with mean 38 years,[127] variance=64 and skewness=10.

  Furthermore, a "true desirability" variable that represents a measure of ability, motivation, productivity and other characteristics that make an applicant desirable for a faculty job was generated as a standardized skew normal with skewness=50 assuming no differences between males and females.

- Hiring model:

  Hiring is assumed to depend on the desirability of the candidate and possibly on gender according to four scenarios: gender neutral (GN), low gender partial (LGP), medium gender partial (MGP) and high gender partial (HGP) (Tables 5.1 and 5.2). A model intercept of -2 was assumed to reflect the log odds of being hired as a male applicant of average desirability.

- Career models:

  Tables 5.1 and 5.2 show model specification and gender coefficients under the four gender partiality scenarios specified.

  Salaries are assumed lognormal with standard deviation 0.2 for the logged salaries. Time in rank for new hires is assumed uniform in $(\frac{1}{12}, \frac{11}{12})$. Annual salary adjustment is assumed lognormal with standard deviation for the logged changes of 0.05.

	Model	
Hiring	$logit\ P(hired_{vt} = 1) =$	$\beta_0 + log(2)\ desirability_{vt} + \beta\ female_v$
Initial rank	$logit\ P(R1_{f0} = 1 \vert R1\ or\ R2) =$ $logit\ P(R2_{f0} = 1 \vert R2\ or\ R3) =$	$-1.5 + 0.25\ MD_{f0} + 0.25\ PhD_{f0} + \beta\ female_f$ $-1.1 + 0.25\ MD_{f0} + 0.2\ PhD_{f0} + \beta\ female_f$
Initial salary	$E[log(ftesal_{df0})] =$	$12 + \beta_{0d} + 0.25\ MD_{df0} + 0.25\ PhD_{df0} + 0.03\ (year_{df0} - 2005)$ $-0.25\ R2_{df0} - 0.50\ R3_{df0} + \beta\ female_{df}$ with $\beta_{0d} \sim N(0, 0.01)$
Departure ($\leq$year 1)	$logit\ P(departure_{f0} = 1) =$	$-2.25 + 0.25\ R2_{f0} + 0.50\ R3_{f0} - 0.25 LftesalC_{f0} + \beta\ female_f$
Promotion	$logit\ P(promoted_{ft} = 1) =$	$-3 + \beta\ female_f$
Change in salary	$E\left[log\left(\frac{ftesal_{dft}}{ftesal_{dft-1}}\right)\right] =$	$0.025 + \beta_{od} + 0.10\ R2.JP_{dft} + 0.10\ R1.JP_{dft} + \beta\ female_{df}$ with $\beta_{0d} \sim N(0, 0.0025)$
Departure (>year 1)	$logP(departure_{ft} = 1) =$	$-2.25 + 0.25\ R2_{ft} + 0.50\ R3_{ft} - 0.25\ LftesalC_{ft} + 0.75\ CSN_{ft}$ $+0.10\ CS1_{ft} - 0.10\ CS2_{ft} - 0.50\ CSM2_{ft} + \beta\ female_f$

$v = 1, ..., V$, with V: number of applicants for a faculty position in a particular year t.
$f = 1, .., F_t$, with F_t: number of faculty members in a particular year t.
$d = 1, ..., D$ with D: number of departments at the university.
$t = 1, ..., T$, with T: number of years in the simulation period and $t = 0$ corresponds to hiring year
Variables:
Ranks: R1=Full, R2=Associate, R3=Assistant
$LftesalC_{ft} = log(ftesal)_{ft}$ centered at $\overline{Lftesal_t}$
Promotion indicators: R1.JP= Full professor just promoted, R2.JP= Associate professor just promoted
Change in salary: CSN= negative change, CS1= change in (0,1]%, CS2= change in (1,2]%, CSM2= change greater than 2%.
Reference category CS0= no change in salary.

Table 5.1: Data generation under known mechanisms of bias: Hiring and Career models

	Gender partiality scenario			
Model	Neutral (GN)	Low (LGP)	Medium (MGP)	High (HGP)
Hiring	0	log(0.85)	log(0.70)	log(0.60)
Initial rank	0	log(0.85)	log(0.70)	log(0.60)
Initial salary	0	-0.06	-0.12	-0.18
Departure ($\leq$year 1)	0	log(1.10)	log(1.20)	log(1.30)
Promotion	0	log(0.85)	log(0.70)	log(0.60)
Change in salary	0	-0.005	-0.01	-0.015
Departure ($>$year 1)	0	log(1.10)	log(1.20)	log(1.30)

Table 5.2: Data generation under known mechanisms of bias: Gender coefficients

The parameters specified create complex gender partiality scenarios as follows:

- For the low gender-partial scenario, as compared to otherwise similar male faculty members, women are: 15% less likely to be hired at all and 15% less likely to be hired at a higher rank, and promoted. They have 10% higher hazard of leaving the institution. Their initial salaries and subsequent annual salary adjustments are 6% and 0.5% lower than men's, respectively.

- For the medium gender-partial scenario, 30% less likely to be hired, to be hired at higher rank and to be promoted. They also have 20% higher hazard of leaving the institution, and their initial salaries and subsequent annual salary adjustments are 12% and 1% lower than men's, respectively.

- For the high gender-partial scenario, 40% less likely to be hired, to be hired at higher rank and to be promoted. They also have 30% higher hazard of leaving the institution, their initial salaries and subsequent annual salary adjustments are 18% and 1.5% lower than men's, respectively.

5.2 Distribution of estimated causal effects

We generated 500 datasets for each gender partiality scenario. For each dataset, we used the simulation-based method to estimate the average causal effects $\overline{HRD}$, $\overline{TRL1}$, $\overline{TRL2}$ and $\overline{RTS}$, with the following parameters:

- condition on observed initial ranks,
- no departures at the end of first year at the institution,
- Weighted counterfactual structure, i.e., set the coefficients of the career models for the gender-neutral scenario ($\hat{\theta}_0$) to be the coefficients of the career models fit for the observed scenario ($\hat{\theta}_1$), with the gender coefficients changed to zero.
- $K = 100$ realizations of the institution.

The weighted counterfactual structure implies that there are no differences for male faculty between the observed institution and its gender-neutral counterpart. Therefore, we present results for the collective of **female faculty** at the institution over the 10 year period under study.

Figure 5.1 shows the distribution of the estimated causal effects for 500 datasets under the four gender partiality scenarios specified (gender neutral, low gender partial, medium gender partial, high gender partial). Table 5.3 shows the corresponding causal effects averaged over the 500 simulated datasets, by specified scenario.

Effect	Gender neutral (GN)	Low gender partial (LGP)	Medium gender partial (MGP)	High gender partial (HGP)
$\overline{RTS}$ [a]	0.07%	4.4%	8.1%	11.9%
$\overline{HRD}$ [b]	1.0	1.1	1.3	1.4
$\overline{TRL2}$ [c]	0.2	-4.0	-6.1	-8.9
$\overline{TRL1}$ [d]	-0.1	-3.2	-7.4	-9.1

[a] Relative Total Salary: % difference in total compensation over 10 years in gender-neutral versus the observed scenario, relative to the observed scenario

[b] Hazard Ratio of Departure: ratio of hazard of departure comparing the observed to the gender neutral scenario

[c] Time at Associate Professor rank Lost: difference in total time at Associate Professor rank between the observed and the gender-neutral scenario, relative to total faculty time

[d] Time at Full Professor rank Lost: difference in total time at Full Professor rank between the observed and the gender-neutral scenario, relative to total faculty time

Table 5.3: Average causal effects over 500 simulated datasets by gender partiality scenario

The gender neutral scenario serves as a control to assess whether the effects are calculated correctly, since in this case both the observed and counterfactual institutions are gender neutral. We can see that the effects estimated in this case are indeed very close to the theoretical values that indicate no departure from gender neutrality (zero for RTS, $TRL2$ and $TRL1$, and 1 for HRD).

The rest of the scenarios reflect increasing differences between the institution and the gender neutral scenario, with increasing gender partiality having a more marked effect on the HRD and RTS than on the time in higher rank lost.

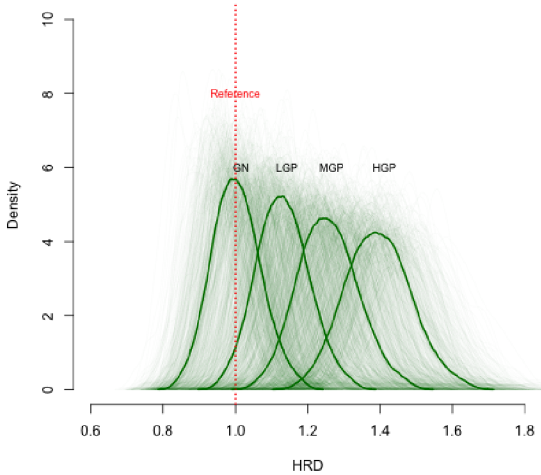
The investment made under gender neutrality on female faculty is 4.4%, 8.1% and 11.9% higher than that made under the low, medium and high gender partiality scenarios specified. This corresponds to 25.8, 36.9 and 51.6 million dollars on average not paid over 10 years to female faculty respectively.

The hazard of departure comparing the simulated datasets to their gender neutral counterparts increases from 10% higher for LGP to 40% higher for HGP.

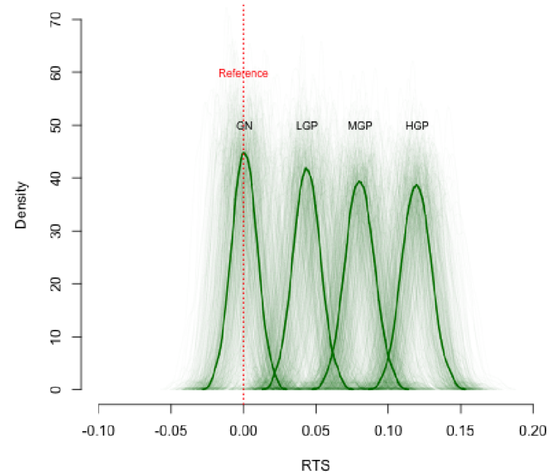
Time in higher rank lost at Associate and Full Professor levels also increases for higher gender partial scenarios, with approximately the same time lost at both levels. For the higher gender-partial scenario for example, female faculty lose 9.1 Full Professor years per 1000 faculty years, which corresponds on average 23 Full Professor years lost to over a period 10 years.

5.3 Distribution of estimated variances of estimated causal effects

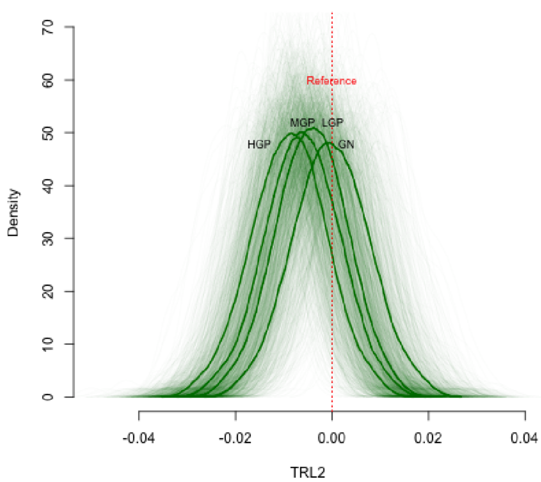
In order to investigate the distribution of the variances of the estimated causal effects, we ran the simulation-based method on 100 simulated institutions under the low gender partial scenario, with 100 iterations and 250 bootstrap replications for each simulated institution. Other scenarios present the same patterns (not shown).



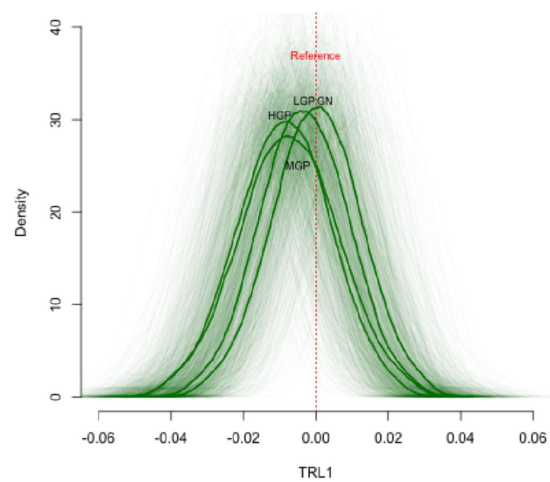
(a) Hazard Ratio of Departure (HRD)



(b) Relative Total Salary



(c) Time in Rank Lost at Associate professor level



(d) Time in Rank Lost at Full professor level

Figure 5.1: Distribution of estimated institution-level causal effects under four bias scenarios: gender neutral (GN), low gender partial (LGP), medium gender partial (MGP) and high gender partial (HGP) for 500 simulated datasets. Each line shown corresponds to a different dataset, with the median density shown as a thicker line. Reference values of no difference between the observed and gender neutral scenarios are shown as a red line.

We estimated the variances of the estimated causal effects in two different ways:

- Method 1: as bootstrap variances of the estimated causal effects calculated over 100 iterations of the simulation-based method.

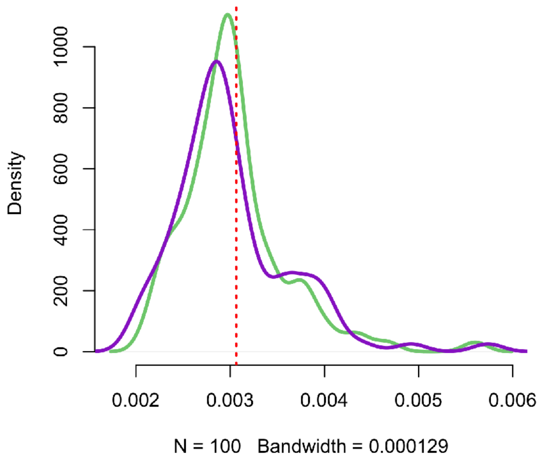
- Method 2: as modeled variances as discussed in section 3.4, i.e, modeling the bootstrap variances using 16 iterations of the method only.

Figure 5.2 shows the distribution of the 100 variances for the 100 simulated institution calculated under both methods.

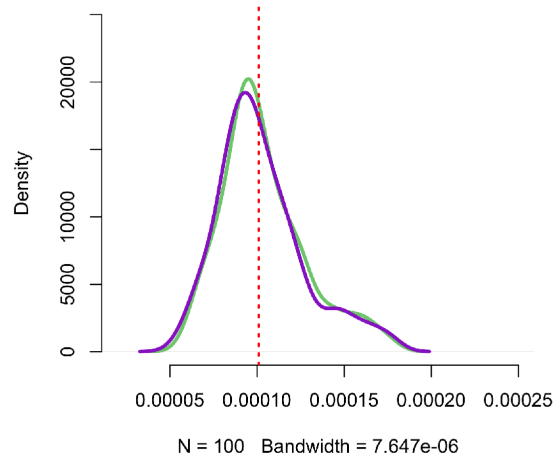
The true variance of the estimated causal effects can be approximated by the variance of the estimated causal effects across the 100 simulated institutions. This is shown as red line in Figure 5.2.

On average, we observe that the estimated variances for $\overline{RTS}$ and $\overline{HRD}$ are within 2.5% of the approximated true variance. The margin of error is larger of $\overline{TRL2}$ and $\overline{TRL1}$, with the estimated variances being approximately 20% higher, which will lead to conservative confidence intervals.

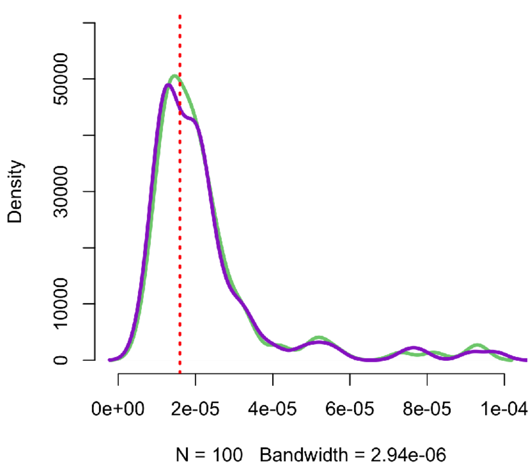
The modeled variances resemble closely the bootstrap variances, so their use would lead to very similar results. Note however that the running time for a single dataset decreases in 85% when using the modeled variances with 16 iterations. This represents a running time of 10 minutes as compared to approximately 40 minutes to run 250 bootstrap iterations on a dataset of approximately 1000 faculty members at the start of the simulation.



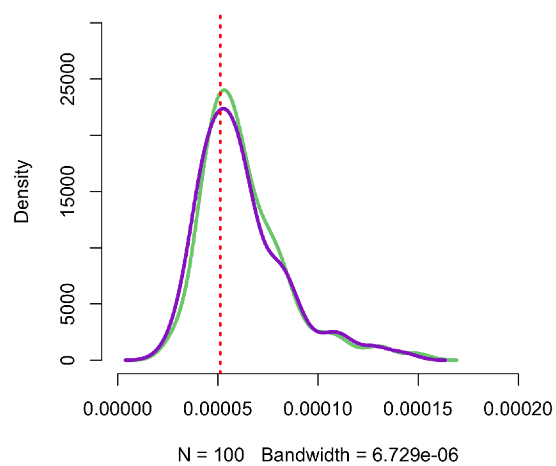
(a) Estimated variance of $\overline{HRD}$



(b) Estimated variance of $\overline{RTS}$



(c) Estimated variance of $\overline{TRL2}$



(d) Estimated variance of $\overline{TRL1}$

Figure 5.2: Distribution of estimated variances of institution-level effects for the low gender partial scenario (LGP) under Method 1 (green) and Method 2 (purple) discussed above. The red line represents approximated true variance of the causal effects.

5.4 Contribution of career steps to estimated causal effects

We used the regression-based method to assess the sources of information for 100 institutions generated under the low gender partiality scenario, using 100 iterations and 250 bootstrap samples.

Each bootstrap sample produces an estimated vector of causal effects $\overline{CE}^{(b)}$ and a vector of gender coefficients $\hat{\theta}_{1G}^{(b)}$. For each simulated institution, and using information from the $b = 1, ..., 250$ bootstrap samples, we regressed the estimated causal effects on the estimated gender coefficients to obtain a t-value and a γ coefficient for each of the career steps. The t-values provide insight as to which career steps contribute more to the estimated causal effects, while the gamma coefficients will provide information on the direction of the relationship.

Tables 5.4 through 5.7 show the average t-values and γ coefficients over the 100 simulated institutions, along with 2.5% and 97.5% quantiles, for each of the causal effects of interest. In this particular example no interactions are significant so only main effects are shown.

The main contributors to the average Relative total compensation ($\overline{RTS}$) are differences in initial salaries between female and male faculty, followed by disparities in annual salary adjustments. In both cases, a larger disparity against female faculty. i.e., lower female initial salaries and annual adjustments with respect to male faculty, increases the value of $\overline{RTS}$.

The average Hazard ratio of departure ($\overline{HRD}$) is mainly determined by differences in the hazards of departure between female and male faculty, followed by disparities in annual salary adjustments and hazard of promotion to Associate Professor. Higher hazards of departure for female faculty, and larger disparities against female faculty in annual salary adjustments and hazard of promotion to Associate professor, produce a higher $\overline{HRD}$

Average Time in higher rank lost at Associate level ($\overline{TRL2}$), is mainly determined by promotions to Full Professor, followed by promotions to Associate Professor. $\overline{TRL2}$ is smaller (more time at Associate Professor level lost) for lower hazards of promotion of female to Full and Associate Professor.

Average Time in higher rank lost at Full level ($\overline{TRL1}$), is mainly determined by promotions to Associate and Full Professor. A higher $\overline{TRL1}$ is associated with lower hazard of promotion of female faculty to Associate level and higher hazard of promotion to Full Professor.

	t-values			γ		
	Mean	Q2.5%	Q97.5%	Mean	Q2.5%	Q97.5%
Initial rank	0.24	-1.79	2.61	0.00	-0.01	0.01
Initial salary	-13.69	-17.32	-10.48	-0.54	-0.68	-0.45
Promotion to Associate	-1.48	-3.18	0.32	-0.00	-0.01	0.00
Promotion to Full	-0.98	-2.84	0.84	-0.00	-0.01	0.00
Annual salary adjustments	-3.11	-5.09	-0.89	-2.52	-4.12	-0.68
Departures	-0.26	-2.44	1.29	-0.00	-0.02	0.01

Table 5.4: Sources of information of $\overline{RTS}$: Average, 2.5% and 97.5% quantiles (Q2.5% and Q97.5%) for t-values and γ coefficients.

	t-values			γ		
	Mean	Q2.5%	Q97.5%	Mean	Q2.5%	Q97.5%
Initial rank	0.17	-1.86	1.99	0.00	-0.01	0.01
Initial salary	0.70	-1.44	3.28	0.03	-0.08	0.16
Promotion to Associate	-4.70	-7.99	-1.63	-0.02	-0.03	-0.01
Promotion to Full	-2.60	-5.67	-0.02	-0.01	-0.02	0.00
Annual salary adjustments	-11.09	-14.69	-8.10	-11.88	-14.97	-9.08
Departures	84.94	74.70	95.12	0.94	0.86	1.02

Table 5.5: Sources of information of $\overline{HRD}$: Average, 2.5% and 97.5% quantiles (Q2.5% and Q97.5%) for t-values and γ coefficients

	t-values			γ		
	Mean	Q2.5%	Q97.5%	Mean	Q2.5%	Q97.5%
Initial rank	-0.75	-3.12	1.95	-0.00	-0.00	0.00
Initial salary	0.20	-1.69	2.13	0.00	-0.01	0.01
Promotion to Associate	5.64	1.95	9.62	0.00	0.00	0.00
Promotion to Full	63.02	42.48	85.53	0.02	0.02	0.03
Annual salary adjustments	-0.31	-2.18	1.86	-0.04	-0.33	0.25
Departures	1.43	-0.56	3.39	0.00	-0.00	0.01

Table 5.6: Sources of information of $\overline{TRL2}$: Average, 2.5% and 97.5% quantiles (Q2.5% and Q97.5%) for t-values and γ coefficients

	t-values			γ		
	Mean	Q2.5%	Q97.5%	Mean	Q2.5%	Q97.5%
Initial rank	1.03	-1.79	3.40	0.00	-0.00	0.01
Initial salary	-0.19	-2.11	1.88	-0.00	-0.02	0.02
Promotion to Associate	48.31	31.26	64.15	0.04	0.03	0.04
Promotion to Full	-34.83	-60.82	-19.92	-0.02	-0.03	-0.01
Annual salary adjustments	-0.13	-2.22	1.82	-0.02	-0.47	0.47
Departures	0.36	-1.63	2.29	0.00	-0.00	0.01

Table 5.7: Sources of information of $\overline{TRL1}$: Average, 2.5% and 97.5% quantiles (Q2.5% and Q97.5%) for t-values and γ coefficients

Chapter 6

Gender equity at University X: institution-level analyses

We applied the simulation-based method to University X, using the following parameters:

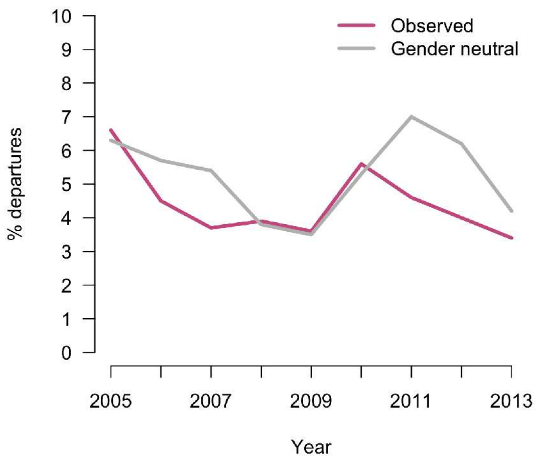
- Time period: 2005-2013,
- Initial ranks observed,
- No departures for newly hired faculty members at the end of first year at the institution,
- Weighted counterfactual structure. This assumes that gender partiality (if it exists) at University X manifests as both under compensation of female faculty and over compensation of male faculty, with the coefficients of the gender-

neutral scenario set to the gender-adjusted coefficients of the observed scenario. Then, the gender-neutral structure for men does not differ from the observed scenario and results presented will be for female faculty only.

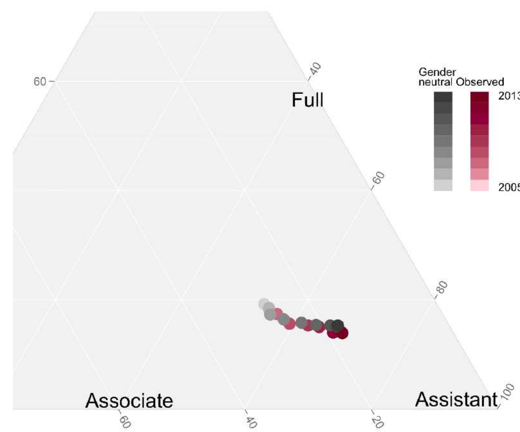
- $K = 100$ realizations of the institution,
- $B = 250$ bootstrap samples for variance calculations.

Figure 6.1 shows comparisons between an observed and simulated gender-neutral realization of University X, chosen at random. For this realization of University X, the simulated observed scenario is very similar to its gender-neutral counterpart. The simulated observed scenario shows slightly lower percentages of female departures over time (6.1a), as well as a slightly lower percentage of female full professors (15.3% versus 13.9% in 2013 respectively) (Figure 6.1b).

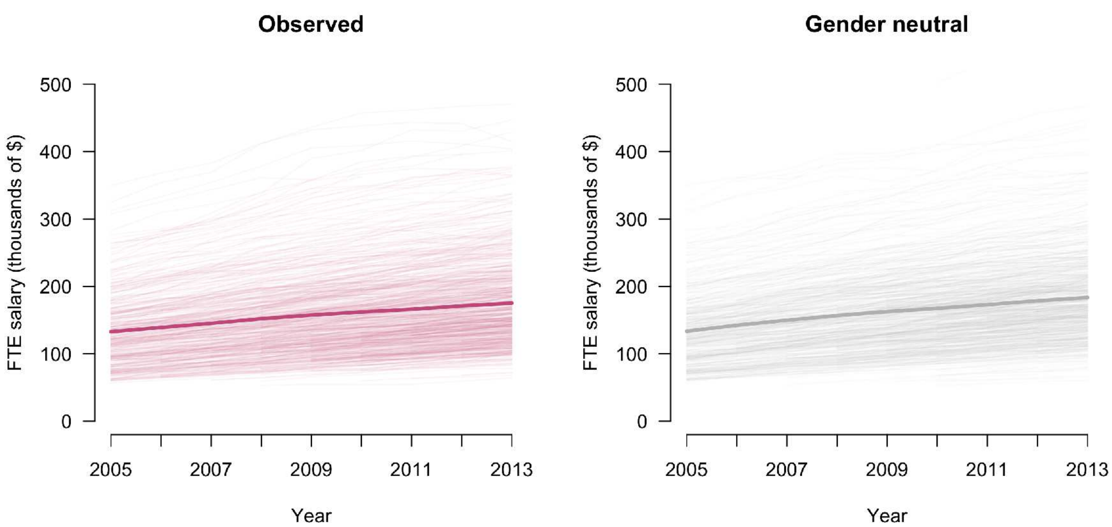
Figure 6.2 shows the estimated average institution-level causal effects of Hazard ratio of departure ($\overline{HRD}$), Relative total compensation ($\overline{RTS}$) and Time in higher rank lost at Associate and Full Professor level ($\overline{TRL2}$ and $\overline{TRL1}$), along with their corresponding distributions over the 100 realizations simulated and corresponding 95% confidence intervals.



(a) % of female departures by scenario over time

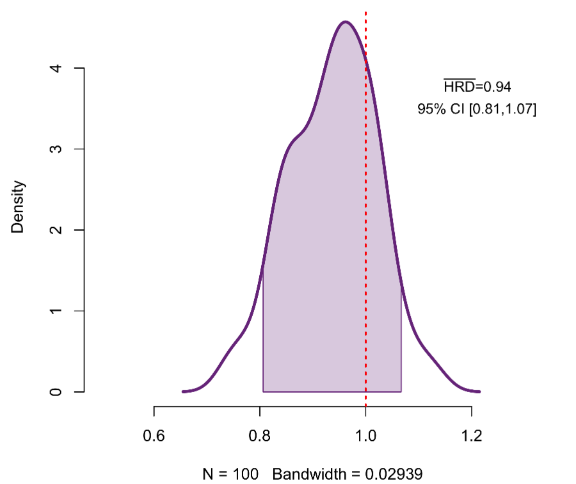


(b) Distribution of faculty ranks by scenario over time. Axes represent % faculty in each rank. For example, for the simulated gender neutral scenario in 2013, 15.3% of women were full professors, 17.6% associate professors and 67.1% assistant professors.

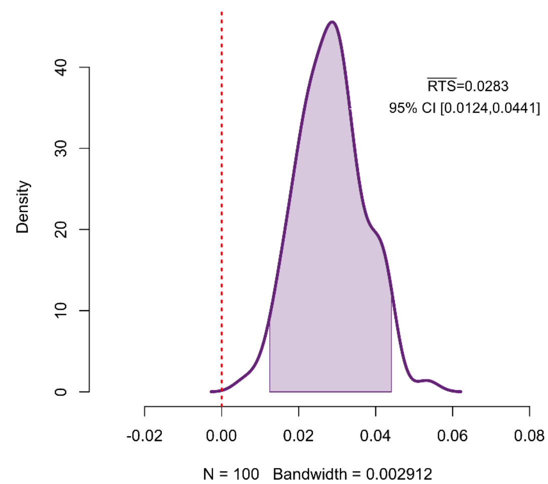


(c) FTE salary by scenario over time

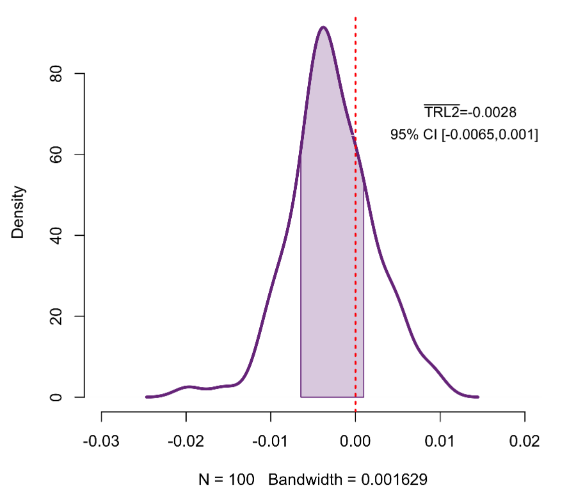
Figure 6.1: University X: example of observed and gender-neutral equity measures 2005-2013 for randomly chosen realization



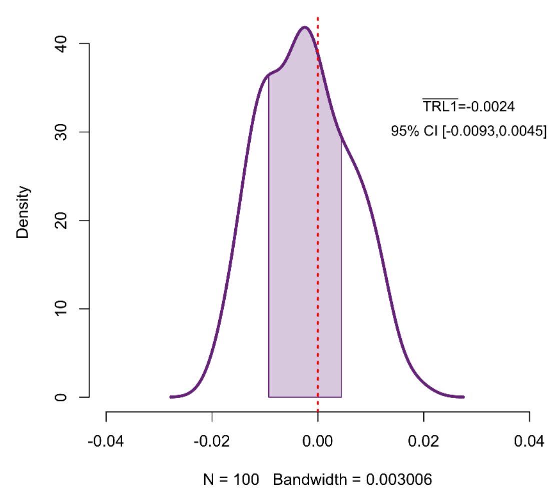
(a) Distribution of Hazard ratio of departure (HRD) over a 100 realizations



(b) Distribution of Relative total compensation (RTS) over a 100 realizations



(c) Distribution of Time at Associate Professor level lost (TRL2) over a 100 realizations



(d) Distribution of Time at Full Professor level lost (TRL1) over 100 realizations

Figure 6.2: University X: Distribution of estimated institution-level causal effects over 100 realizations. Shaded areas correspond to 95% confidence intervals. Red line represents reference value of no difference between observed and gender neutral scenarios.

There are no significant departures from gender neutrality in the case of departures and time at Full or Associate professor level. This means that female faculty depart University X as expected if University X were gender neutral, and female faculty do not lose any time in higher ranks. However, the percentage of money spent on female faculty during 2005-2013 is 2.8% (95% CI [1.2%,4.4%]) higher under gender neutrality.

The biggest contributor to the departures from gender neutrality in total compensation over the 9 study years correspond to differences in initial salaries (Table 6.1), with a t-value of -17.82, followed by differences in the annual salary adjustments (t-value=-5.78) and differences in the promotion to associate professor (t-value=-3.01). The rest of the the steps of the career do not have a significant contribution to the average Relative total compensation ($\overline{RTS}$). There were no significant interaction between any of the career steps.

Negative $\hat{\beta}$ coefficients for initial salary, annual salary adjustments and promotion to Associate Professor indicate that larger gender disparities against female faculty in these steps contribute to making $\overline{RTS}$ larger, hence to larger departures from gender neutrality. On average:

- making initial female faculty salaries one extra percent point lower than male's is associated with an increase of 0.005 in $\overline{RTS}$,
- lowering the annual salary adjustments for female faculty 0.1% with respect to males increases the $\overline{RTS}$ by 0.00332,

- decreasing the hazard of promotion to Associate Professor 10 extra percent points, increases $\overline{RTS}$ in 0.00095.

	$\hat{\beta}^{*}$	$\widehat{SE}(\hat{\beta})$	t value	p-value
Initial salary	-0.4970	0.0279	-17.82	$<$0.001
Annual salary adjustment	-3.3185	0.5739	-5.78	$<$0.001
Promotion to Associate	-0.0095	0.0032	-3.01	0.0029
Promotion to Full	-0.0031	0.0021	-1.48	0.14
Departures	0.0014	0.0037	0.38	0.70
Initial rank	0.0008	0.0023	0.35	0.73
Departure 2005[a]	0.0004	0.0013	0.28	0.78

$^{*}\beta$ corresponds to average change in $\overline{RTS}$ per unit change in the gender coefficient for each of the career models after adjusting for the rest

[a] corresponds to an extra model for the hazard of departure for existing faculty in 2005

Table 6.1: University X: contribution of career models to the average Relative total compensation $\overline{RTS}$

Chapter 7

Discussion

In this dissertation we propose novel methods to estimate higher-level causal effects in complex systems, as well as assess the influence of individual components of the system on the causal estimates. The methods proposed (1) account for individuals nested within a larger cluster where the intervention or condition of interest is applied to all the individuals in a complex way and (2) address the specific situation where the aggregation level of interest corresponds to the cluster of all individuals, so only one observation of the outcomes of interest is available (a sample size of 1). This methodology complements the existing literature on the estimation of causal effects when measurements are available for multiple units from the target population.

These methods were motivated by the study of gender equity in academia for a single institution. Previous work has focused on the individual faculty member as a study unit, and in most applications, on a single academic reward or representation outcome. The idea behind this methodology to assess gender equity is that the understanding of each step of the academic career will give us insight into the behavior of the university as a whole. However, since academic institutions are complex systems, by reducing their study to a single academic reward, we are ignoring the influences of and the interactions between the academic rewards.

We defined the intervention of interest as "gender partiality" in contrast to "gender neutrality". The first term implies institutional practices that result in unequal access to opportunities for female and male faculty, and which reflect on key academic rewards and representation outcomes such as lower initial salaries and ranks for women, smaller annual adjustments in salary and higher rates of departure from the institution. The use of "gender partiality" as the intervention or exposure of interest instead of "gender" shifts the perspective of measurement from the individual to the institution and attempts to reflect that the institution's perception of gender, and not gender itself, is the potential reason for disparities in academic rewards. This consideration is central to this dissertation since the use of this perspective overcomes criticisms exposed in the literature regarding the nature of gender as a causal variable.

We consider then the **institution** as the unit of analysis, a system made up of the individual faculty member's careers in seven steps: hiring, initial salary, initial rank, promotion to Associate Professor, promotion to Full Professor, annual adjustments in salary and departure from the institution. We focused on single-institution studies, and defined the following institution-level outcomes for the collective of all faculty members at the institution, over a period of T years:

- Hazard of departure from the institution,
- Time in higher rank (Associate or Full Professor level),
- Total compensation.

From a causal inference perspective, we would like to have information for any single institution under two scenarios: the observed scenario (potentially gender partial) and a gender neutral scenario. In other words, to uncover the effect of potentially gender partial practices in academic rewards and representation outcomes, we would need to compare the observed institution to the exact same institution, under the assumption of gender neutrality. This is the fundamental problem of causal inference: we can only observe one of these potential outcomes.

What we then propose is a method that allows us to simulate the same institution under both the observed and gender neutral scenarios, based on the fact that the institution-level outcomes are functions of individual-level measurements, where we do have multiple units. These methods allow us to:

1. Estimate institution-level causal effects,
2. Estimate the variance of the institution-level causal effects,
3. Determine the influence of individual components on the estimated causal effects.

The simulation-based method estimates the average effects in 5 steps:

1. Develop a model for the observed system,

2. Estimate the model from the observed data,

3. Posit a model for the counterfactual system,

4. Simulate K realizations of the system under both scenarios. In each realization estimate the institution-level causal effects,

5. Average over all realizations to obtain the average causal effects.

In order to actually assign causal interpretations to the effects calculated, the assumptions of correct model specification and ergodicity (as defined by Marshall and Galea (2014)[21]) need to hold.

The first assumption, correct model specification, implies that models reflect the mechanisms at place in the institution and all relevant variables are included in the analyses.[21] This also means that the models need to be well-calibrated and able to reproduce the observed data patterns.[21]

The models used here represent but one possible system behavior of an academic institution and specific applications need to be tailored to the university under study. The method proposed can easily accommodate different mechanisms by updating the nested career models to the necessary functional form and relevant variables for a particular institution. Examples of this situation include the use of productivity measures in the analyses or assuming different rewards by academic rank.

The assumption of ergodicity means that the average of the outcomes across realizations are well-defined, or that the system is stable,[21] so the causal effects converge to a fixed value as the number of realizations simulated increases. This assumption can be investigated using the procedures outlined by Grazzini (2012).[132]

Results for simulated data showed that the simulation-based method correctly captures increasing levels of gender partiality for the institution-level causal effects. In the case of the gender neutral scenario simulation, observed non-significant departures from the theoretical values of zero for $\overline{RTS}$, $\overline{TRL2}$ and $\overline{TRL1}$, and 1 for $\overline{HRD}$ (Table 5.3) might reflect two situations (1) the number of datasets generated is not large enough to reach the theoretical average and (2) since the existing faculty distribution is simulated under labor market conditions and assumed gender neutral to simulate the careers for both observed and counterfactual simulations, spillover gender partiality effects may still exist after 10 years. This could be assessed by separating the causal effect in two parts: one attributable to new hires, and another attributable to existing faculty members at the start of the simulation.

The variance of the estimated causal effects is estimated using bootstrapping procedures. We further propose the use of linear modeling to reduce the number of iterations needed during bootstrap to estimate the variances. This method, along with implemented code in R and C++ through Rcpp reduces running time by approximately 75%. Future research on this topic will include finding a combination of the number of iterations and bootstrap samples that allow for optimal estimates of the variances in a short running time.

When we applied our methods to University X data, we did not find significant deviations from gender neutrality in terms of departures of female faculty from the University, with an average hazard of departure 6% lower than the hypothetical gender neutral institution ($\overline{HRD}$=0.94 95% CI [0.81,1.07]). There are also no statistically significant losses of time in higher rank at Associate or Full Professor levels. We estimate that 2.8 Full Professor years (95% CI [1.0,6.5]) and 2.4 Associate Professor years (95% CI [4.5, 9.3]) are lost per 1000 female faculty years.

However, University X does not appear to be gender neutral with regard to the total female faculty compensation over 9 years. When University X is made gender neutral by setting its estimated gender coefficients to 0.0, the total female salaries are 3% higher representing 17.5$ million over the last nine years (95% CI [7.7,27.3].

A major limitation of the analysis above is the lack of productivity measures, which are not currently available for University X. Productivity is a known determinant of academic rewards. The models specified for University X might be incorrectly specified if productivity is also linked to gender.

An advantage of the simulation-based method is that even though we may not have information for all potential confounders (such as productivity), we can hypothesize their behaviors and include them in the nested career models to assess their influence on the estimated causal effects. This would constitute a sensitivity analysis that would allow us to decide whether the collection of further information is needed to assess gender equity at the institution.

The choice of the appropriate gender neutral counterfactual has been controversial in the decomposition technique literature. However, just as in that case, the simulation-based method allows for comparison of different counterfactual structures.

The current method is set to handle binary exposures, however it can be naturally extended to discrete exposures with more than 2 categories or to continuous exposures. For example, we can assess race partiality at an institutional-level by comparing the outcomes for White, Asian and Underrepresented Minorities to a counterfactual where faculty members have access to the same opportunities regardless of race.

Academic institutions are but one example of complex systems in which we might be interested in assessing higher-level effects, and the methods presented in this dissertation are applicable to a wide variety of topics. The following are examples of additional applications:

- Study of the effect a new law would have on the total revenue of a hospital system as a function of the practice of each of its doctors,

- Investigation of the effect a new technological development would have on the time from manufacturer to consumer of a drug, as a function of each of the development stages.

Despite the differences that still exist in rewards and representation in academia between female and male faculty, there has been progress in the last decade in the achievement of a more gender-neutral structure. We hope this dissertation will be one more tool for academic institutions to provide all their faculty, regardless of gender, with the same opportunities for success.

Chapter 8

Appendix

8.1 Comments on the magnitude and shape of contrast variances (Chapter 4)

The magnitude of the variances of the longitudinal contrasts $Var(\bar{y}_{L1i} - \bar{y}_{L0i})$ is largely determined by the variance of the random slope on treatment σ_I^2 and, in a smaller level, by the measurement error σ_e^2 and the number of people per cluster J. The larger the σ_I^2, σ_e^2 and the smaller J, the larger the variance of the longitudinal contrast.

The shape of the variances across clusters (in order of assignment to intervention) is determined by the total number of measurements K and the number of measurements before and after the cluster is assigned to the intervention (K_{0i} and K_{1i}). The greater the imbalance in the number of observations in the control / intervention periods, the larger $Var(\bar{y}_{L1i} - \bar{y}_{L0i})$. For a fixed number of people in each cluster, the variance of the contrasts is symmetrical, with lower magnitudes near the middle of the assignment-to-intervention period. The variance of the cluster-level random intercept σ_C^2 does not play a role in the variance of the longitudinal contrasts.

The variance of the cross-sectional contrasts $Var(\bar{y}_{C1t} - \bar{y}_{C0t})$ is largely influenced by the number of clusters assigned to the intervention at a particular time point (N_{1t}). The larger the imbalance between the number of clusters assigned to control/intervention, the larger $Var(\bar{y}_{C1t} - \bar{y}_{C0t})$. σ_I^2 determines the shape of $Var(\bar{y}_{C1t} - \bar{y}_{C0t})$ across time. For $\sigma_I^2 = 0$, the variances are symmetric, while for $\sigma_I^2 > 0$, the variances of the earlier contrasts is larger than the latter. The magnitude of variances depend on σ_I^2, σ_C^2, σ_P^2, σ_e^2 and J, with larger variances of random effects and smaller number of people per cluster resulting in larger contrast variances.

For fixed number of measurements K, number of clusters N and assignment-to-intervention period, the variance of the longitudinal contrasts are always smaller than the variances of the cross-sectional contrasts when $\sigma_I^2 = 0$, regardless of J, σ_e^2e and σ_C^2. When $\sigma_I^2 > 0$ and $\sigma_C^2 = 0$, the variances of the longitudinal contrasts are generally larger than the variances of most of the cross-sectional contrasts, with the exception of the earlier ones. However, as the magnitude of the cross-sectional variances increases with σ_C^2 (and the variances of the longitudinal do not), the bigger the difference between the variances of cross-sectional and longitudinal contrasts.

8.2 Parameters considered to generate stepped wedge data (Chapter 4)

We generated data for a gaussian response as measured in a stepped wedge design following model 4.1, for combinations of several values of the parameters J, K, N, β_1, σ_C^2, σ_I^2, as follows:

- Linear trend in time and seasonality: both absent, trend present but seasonality absent, both trend and seasonality present.
- Cluster size J: number of individuals per cluster: 5, 10 or 20.
- Cluster size type: fixed or random number of individuals per cluster. If random, it is generated as a random draw from a Poisson distribution with mean J.
- Number of measurements K: 4, 12, 24 months.
- Number of clusters (N): 1/4 or 1/2 the number of measurements.
- Treatment effect: $\beta_1 = -0.1$ or $\beta_1 = -0.5$.
- Cluster-level variance: $\sigma_C^2 = 0$ or $\sigma_C^2 = 0.5$.
- Random slope on treatment effects: yes or no.

Bibliography

[1] N. R. C. (U.S.), "Gender differences at critical transitions in the careers of science, engineering, and mathematics faculty," National Academy Press, Tech. Rep., 2010.

[2] L. W. Perna, "Sex differences in faculty tenure and promotion: The contribution of family ties," *Research in Higher Education*, vol. 46, no. 3, pp. pp. 277–307, 2005.

[3] D. A. Barbezat and J. W. Hughes, "Salary structure effects and the gender pay gap in academia," *Research in Higher Education*, vol. 46, no. 6, pp. 621–640, 2005.

[4] L. Fox-Cardamone, "Inequity in the academy: A case study of factors influencing promotion and compensation in american universities." in *Forum on Public Policy Online*, vol. 2010, no. 5. ERIC, 2010.

[5] J. H. Miller and S. E. Page, *Complex Adaptive Systems: An Introduction*

to Computational Models of Social Life (Princeton Studies in Complexity). Princeton, NJ, USA: Princeton University Press, 2007.

[6] L. W. Perna, "Studying faculty salary equity: A review of theoretical and methodological approaches," in *Higher education: Handbook of theory and research*. Springer, 2003, pp. 323–388.

[7] B. N. Torgrimson and C. T. Minson, "Sex and gender: what is the difference?" *Journal of Applied Physiology*, vol. 99, no. 3, pp. 785–787, 2005. [Online]. Available: http://jap.physiology.org/content/99/3/785

[8] S. Borna and G. White, ""sex" and "gender": Two confused and confusing concepts in the "women in corporate management" literature," *Journal of Business Ethics*, vol. 47, no. 2, pp. 89–99, 2003. [Online]. Available: http://www.jstor.org/stable/25075129

[9] R. K. Unger and M. Crawford, "Commentary: Sex and gender: The troubled relationship between terms and concepts," *Psychological Science*, vol. 4, no. 2, pp. 122–124, 1993. [Online]. Available: http://www.jstor.org/stable/40062521

[10] P. W. Holland, "Causation and race," *White logic, White methods*, pp. 93–109, 2008.

[11] M. Hernan, "A definition of causal effect for epidemiological research," *Journal of Epidemiology and Community Health*, vol. 58, no. 4, pp. 265–271, 2004.

[12] P. W. Holland, "Statistics and causal inference," *Journal of the American statistical Association*, vol. 81, no. 396, pp. 945–960, 1986.

[13] D. B. Rubin, "Statistics and causal inference: Comment: Which ifs have causal answers," *Journal of the American Statistical Association*, vol. 81, no. 396, pp. 961–962, 1986.

[14] J. S. Kaufman and R. S. Cooper, "Seeking causal explanations in social epidemiology," *American journal of epidemiology*, vol. 150, no. 2, pp. 113–120, 1999.

[15] C. Glymour, "Statistics and causal inference: Comment: Statistics and metaphysics l nu®."

[16] D. J. Greiner and D. B. Rubin, "Causal effects of perceived immutable characteristics," *Review of Economics and Statistics*, vol. 93, no. 3, pp. 775–785, 2011.

[17] M. Bertrand and S. Mullainathan, "Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination," National Bureau of Economic Research, Tech. Rep., 2003.

[18] C. Goldin and C. Rouse, "Orchestrating impartiality: The impact of "blind" auditions on female musicians," National bureau of economic research, Tech. Rep., 1997.

[19] B. Pavlic, L. Ruprecht, and S. Sam-Vargas, "Gender equality and equity: A summary review of unesco's accomplishments since the fourth world conference on women (beijing 1995)," Unit for the Promotion of the Status of Women and Gender Equality, UNESCO, Tech. Rep., 1995.

[20] C. Guzmán-Valenzuela and R. Cortés, "Constructing academic identities: between the academic career and the academic trajectory." Society for Research into Higher Education (SRHE) Annual Research Conference, 2013.

[21] B. D. Marshall and S. Galea, "Formalizing the role of agent-based modeling in causal inference and epidemiology," *American journal of epidemiology*, p. kwu274, 2014.

[22] J. S. Long, "From scarcity to visibility: Gender differences in the careers of doctoral scientists and engineers," National Academy Press, Tech. Rep., 2001.

[23] A. C. Ellis, "Preliminary report of committee w, on status of women in college and university faculties," *Bulletin of the American Association of University Professors*, vol. 7, no. 6, pp. 21–32, 1921.

[24] N. S. Foundation, "Women as a percentage of full-time, full professors with science, engineering, and health doctorates, by employing institution: 1993–2013," [Accessed 11-March-2016]. [Online]. Available: http://www.nsf.gov/statistics/2015/nsf15311/digest/theme7.cfm

[25] D. L. Stacy and C. L. Holland Jr, "Legal and statistical problems in litigating sex discrimination claims in higher education." *Journal of College and University Law*, vol. 11, no. 2, pp. 107–77, 1984.

[26] N. S. Foundation, "Doctorate recipients, by sex and broad field of study: Selected years, 1984–2014," [Accessed 11-March-2016]. [Online]. Available: http://www.nsf.gov/statistics/2016/nsf16300/data/tab14.pdf

[27] AAMC, "Total graduates by u.s. medical school and sex, 2010-2011 through 2014-2015," [Accessed 29-September-2015]. [Online]. Available: https://www.aamc.org/download/321532/data/factstableb2-2.pdf

[28] H. Shen, "Inequality quantified: Mind the gender gap." *Nature*, vol. 495, no. 7439, pp. 22 – 24, 2013. [Online]. Available: http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=85914672&site=ehost-live&scope=site

[29] R. E. Steinpreis and D. Anders, Katie A.and Ritzke, "The impact of gender on the review of the curricula vitae of job applicants and tenure candidates: A national empirical study," *Sex Roles*, vol. 41, no. 7, pp. 509–528. [Online]. Available: http://dx.doi.org/10.1023/A:1018839203698

[30] N. R. Council, "To recruit and advance: Women students and faculty in u.s. science and engineering," National Academies Press, Tech. Rep., 2006.

[31] R. F. T. Stephen J. Ceci, Wendy M. Williams, "Understanding current

causes of women's underrepresentation in science," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 8, pp. 3157–3162, 2011. [Online]. Available: http://www.jstor.org/stable/41060898

[32] v. . . n. . . p. . . y. . . d. . p. U. . h. e. . h. j. . P. Wendy M. Williams and Stephen J. Ceci, title = National hiring experiments reveal 2:1 faculty preference for women on STEM tenure track.

[33] Y. J. Xu, "Gender disparity in stem disciplines: A study of faculty attrition and turnover intentions," *Research in Higher Education*, vol. 49, no. 7, pp. 607–624, 2008. [Online]. Available: http://dx.doi.org/10.1007/s11162-008-9097-4

[34] B. CL and L. V, "Addressing disparities in academic medicine: Moving forward," *JAMA*, vol. 314, no. 11, pp. 1139–1141, 2015. [Online]. Available: +http://dx.doi.org/10.1001/jama.2015.10664

[35] J. M. Sheltzer and J. C. Smith, "Elite male faculty in the life sciences employ fewer women," *Proceedings of the National Academy of Sciences*, vol. 111, no. 28, pp. 10 107–10 112, 2014. [Online]. Available: http://www.pnas.org/content/111/28/10107.abstract

[36] J. Loeb and M. Ferber, "Sex as predictive of salary and status on a university faculty," *Journal of educational measurement*, vol. 8, no. 4, pp. 235–244, 1971.

[37] R. K. Toutkoushian, *Analyzing equity in faculty compensation.* John Wiley & Sons, 2012.

[38] Y. Weiss and L. Lillard, "Output variability, academic labor contracts, and waiting times for promotion," *Research in Labor Econonomics*, vol. 5, pp. 157–188, 1982.

[39] R. M. J. Scott Long, Paul D. Allison, "Rank advancement in academic careers: Sex differences and the effects of productivity," *American Sociological Review*, vol. 58, no. 5, pp. 703–722, 1993. [Online]. Available: http://www.jstor.org/stable/2096282

[40] K. J. H. Donna K. Ginther, "Gender differences in salary and promotion in the humanities," *The American Economic Review*, vol. 89, no. 2, pp. 397–402, 1999. [Online]. Available: http://www.jstor.org/stable/117143

[41] J. C. Kissane, "Status of comparable worth: Spaulding v. university of washington, the," *Del. J. Corp. L.*, vol. 10, p. 119, 1985.

[42] L. N. Gasaway, "Comparable worth: A post-gunther overview," *Geo. LJ*, vol. 69, p. 1123, 1980.

[43] U. E. E. O. Commission, "Equal pay act of 1963 and lilly ledbetter fair pay act of 2009," [Accessed 15-March-2016]. [Online]. Available: http://www.eeoc.gov/eeoc/publications/brochure-equal_pay_and_ledbetter_act.cfm

[44] D. of Science Resources Statistics, “Gender differences in the careers of academic scientists and engineers: A literarure review,” National Science Foundation, Tech. Rep., 2003.

[45] M. K. Clayton, “A statistical investigation of gender equity in salaries of faculty at the university of minnesota,” University of Wisconsin-Madison, Tech. Rep., 2006.

[46] O. of the Vice Provost for the Faculty, “Report on the us berkeley faculty salary equity study,” University of California, Berkeley, Tech. Rep., 2015.

[47] ——, “Gender equity assessment: Faculty compensation and headcounts,” University of Texas at Austin, Tech. Rep., 2013.

[48] T. F. on the Status of Women at IUPUI, “Detailed report: Task force on the status of women at iupui,” University of Indiana - Purdue, Tech. Rep., 2014.

[49] A. S. Ash, P. L. Carr, R. Goldstein, and R. H. Friedman, “Compensation and advancement of women in academic medicine: Is there equity?” *Annals of Internal Medicine*, vol. 141, no. 3, pp. 205–212, 2004. [Online]. Available: +http://dx.doi.org/10.7326/0003-4819-141-3-200408030-00009

[50] A. Jena, D. Khullar, O. Ho, A. Olenski, and D. Blumenthal, “Sex differences in academic rank in us medical schools in 2014,” *JAMA*, vol. 314, no. 11, pp. 1149–1158, 2015. [Online]. Available: +http://dx.doi.org/10.1001/jama.2015.10680

[51] P. Ohm, "Broken promises of privacy: Responding to the surprising failure of anonymization," *UCLA Law Review*, vol. 57, p. 1701, 2010.

[52] K. Purdam and M. Elliot, "A case study of the impact of statistical disclosure control on data quality in the individual uk samples of anonymised records," *Environment and Planning A*, vol. 39, no. 5, p. 1101, 2007.

[53] B. Nowok, G. M. Raab, and C. Dibben, "synthpop: Bespoke creation of synthetic data in r."

[54] ——, *synthpop: Generating Synthetic Versions of Sensitive Microdata for Statistical Disclosure Control*, 2015, r package version 1.1-0. [Online]. Available: http://CRAN.R-project.org/package=synthpop

[55] D. A. Barbezat, "History of pay equity studies," *New Directions for Institutional Research*, vol. 2002, no. 115, pp. 9–40, 2002. [Online]. Available: http://dx.doi.org/10.1002/ir.52

[56] R. L. Marschke, "Gender pay equity among university faculty: Testing segregation, human capital, discrimination, and demographic inertia," Ph.D. dissertation, University of Colorado, 1997.

[57] L. Braskamp, J. Muffo, and I. Langston, "Determining salary equity: Policies, procedures, and problems," *The Journal of Higher Education*, vol. 49, no. 3, pp. 231–246, 1978. [Online]. Available: http://www.jstor.org/stable/1981113

[58] N. A. Bereman and J. A. Scott, "Using the compa-ratio to detect gender bias in faculty salaries," *The Journal of Higher Education*, pp. 556–569, 1991.

[59] R. D. Raymond, M. L. Sesnowitz, and D. R. Williams, "Further evidence on gender and academic rank," *The Quarterly Review of Economics and Finance*, vol. 33, no. 2, pp. 197–215, 1993.

[60] M. A. Ferber and C. A. Green, "Traditional or reverse sex discrimination-a case study of a large public university," *Indus. & Lab. Rel. Rev.*, vol. 35, p. 550, 1981.

[61] M. R. Ransom and S. B. Megdal, "Sex differences in the academic labor market in the affirmative action era," *Economics of Education Review*, vol. 12, no. 1, pp. 21–43, 1993.

[62] S. Kahn, "Gender differences in academic career paths of economists," *The American Economic Review*, vol. 83, no. 2, pp. 52–56, 1993. [Online]. Available: http://www.jstor.org/stable/2117639

[63] J. W. Loeb, "Hierarchical linear modeling in salary equity studies," *New Directions for Institutional Research*, vol. 2003, no. 117, pp. 69–96, 2003. [Online]. Available: http://dx.doi.org/10.1002/ir.69

[64] P. D. Umbach, "Gender equity in college faculty pay: A cross-classified random effects model examining the impact of human capital, academic disciplines,

and institutions," in *Annual Meeting of the American Educational Research Association, New York, NY, March*, 2008, pp. 24–28.

[65] P. Umbach, "Gender equity in the academic labor market: An analysis of academic disciplines," *Research in Higher Education*, vol. 48, no. 2, pp. 169–192, 2007.

[66] D. A. Katz, "Faculty salaries, promotions, and productivity at a large university," *The American Economic Review*, vol. 63, no. 3, pp. 469–477, 1973.

[67] S. Farber, "The earnings and promotion of women faculty: Comment," *The American Economic Review*, vol. 67, no. 2, pp. 199–206, 1977.

[68] B. T. Hirsch and K. Leppel, "Sex discrimination in faculty salaries: Evidence from a historically women's university," *The American Economic Review*, vol. 72, no. 4, pp. 829–835, 1982.

[69] P. S. Tolbert, "Organizations and inequality: Sources of earnings differences between male and female faculty," *Sociology of Education*, vol. 59, no. 4, pp. 227–236, 1986. [Online]. Available: http://www.jstor.org/stable/2112349

[70] J. P. Formby, W. D. Gunther, and R. Sakano, "Entry level salaries of academic economists: Does gender or age matter?" *Economic Inquiry*, vol. 31, no. 1, pp. 128–138, 1993.

[71] J. Monks and M. Robinson, "Gender and racial earnings differentials in academic labor markets," *Economic Inquiry*, vol. 38, no. 4, pp. 662–671, 2000.

[72] M. Binder, K. Krause, J. Chermak, J. Thacher, and J. Gilroy, "Same work, different pay? evidence from a us public university," *Feminist Economics*, vol. 16, no. 4, pp. 105–135, 2010. [Online]. Available: http://dx.doi.org/10.1080/13545701.2010.530605

[73] R. K. Toutkoushian, M. L. Bellas, and J. V. Moore, "The interaction effects of gender, race, and marital status on faculty salaries," *The Journal of Higher Education*, vol. 78, no. 5, pp. 572–601, 2007.

[74] M. L. Bellas, "Faculty salaries: Still a cost of being female?" *Social Science Quarterly*, vol. 74, no. 1, pp. 62–75, 1993.

[75] R. K. Toukoushian and E. P. Hoffman, "Alternatives for measuring the unexplained wage gap," *New Directions for Institutional Research*, vol. 2002, no. 115, pp. 71–90, 2002. [Online]. Available: http://dx.doi.org/10.1002/ir.54

[76] M. Fox, "Location, sex-typing, and salary among academics," *Work and Occupations: An International Sociological Journal*, vol. 2, no. 12, pp. 186–205, 1985.

[77] G. V. Barrett, R. A. Alexander, M. N. Anesgart, and D. Doverspike, "Frequently encountered problems in the application of regression analysis

to the investigation of sex discrimination in salaries," *Public Personnel Management*, vol. 15, no. 2, pp. 143–157, 1986. [Online]. Available: http://ppm.sagepub.com/content/15/2/143.abstract

[78] D. A. Conway and H. V. Roberts, "Reverse regression, fairness, and employment discrimination," *Journal of Business ans Economic Statistics*, vol. 1, no. 1, pp. 75–85, 1983. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/07350015.1983.10509326

[79] A. S. Goldberger, "Reverse regression and salary discrimination," *The Journal of Human Resources*, vol. 19, no. 3, pp. 293–318, 1984. [Online]. Available: http://www.jstor.org/stable/145875

[80] W. H. Greene, "Reverse regression: The algebra of discrimination," *Journal of Business and Economic Statistics*, vol. 2, no. 2, pp. 117–120, 1984. [Online]. Available: http://www.jstor.org/stable/1391254

[81] M. Nzeukou, "Detecting gender salary inequity in the presence of within gender inequities," Ph.D. dissertation, University of Arizona, 2006.

[82] C. A. Green and M. A. Ferber, "Employment discrimination: An empirical test of forward versus reverse regression," *Journal of Human Resources*, pp. 557–569, 1984.

[83] D. E. Bloom and M. R. Killingsworth, "Pay discrimination research and litiga-

tion: The use of regression," *Industrial Relations: A Journal of Economy and Society*, vol. 21, no. 3, pp. 318–339, 1982.

[84] R. Oaxaca, "Male-female wage differentials in urban labor markets," *International economic review*, vol. 14, no. 3, pp. 693–709, 1973.

[85] A. S. Blinder, "Wage discrimination: reduced form and structural estimates," *Journal of Human resources*, pp. 436–455, 1973.

[86] R. L. Oaxaca and M. Ransom, "Calculation of approximate variances for wage decomposition differentials," *Journal of Economic and Social Measurement*, vol. 24, no. 1, pp. 55–61, 1998.

[87] B. Jann, "The blinder-oaxaca decomposition for linear regression models," *The Stata Journal*, vol. 8, no. 4, pp. 453–479, 2008.

[88] R. W. Fairlie, "The absence of the african-american owned business: An analysis of the dynamics of self-employment," *Journal of Labor Economics*, vol. 17, no. 1, pp. 80–108, 1999. [Online]. Available: http://www.jstor.org/stable/10.1086/209914

[89] R. Fairlie, "An extension of the blinder-oaxaca decomposition technique to logit and probit models," 2014. [Online]. Available: http://www.escholarship.org/uc/item/7j86n5q3

[90] D. Neumark, “Employers’ discriminatory behavior and the estimation of wage discrimination,” *Journal of Human resources*, pp. 279–295, 1988.

[91] J. Cotton, “On the decomposition of wage differentials,” *The review of economics and statistics*, pp. 236–243, 1988.

[92] B. C. Buzan and T. L. Hunt, “Evaluating faculty performance under the equal pay for equal work doctrine,” *Research in Higher Education*, vol. 5, no. 2, pp. 113–123, 1976. [Online]. Available: http://www.jstor.org/stable/40195053

[93] D. A. Barbezat, “Salary differentials by sex in the academic labor market,” *The Journal of Human Resources*, vol. 22, no. 3, pp. 422–428, 1987.

[94] ——, “Salary differentials or sex discrimination? evidence from the academic labor market,” *Population Research and Policy Review*, vol. 6, no. 1, pp. 69–84, 1987. [Online]. Available: http://www.jstor.org/stable/40230002

[95] Y. Haberfeld and Y. Shenhav, “Are women and blacks closing the gap? salary discrimination in american science during the 1970s and 1980s,” *Industrial and Labor Relations Review*, vol. 44, no. 1, pp. 68–82, 1990. [Online]. Available: http://www.jstor.org/stable/2523430

[96] J. Ashraf, “The influence of gender on faculty salaries in the united states, 1969-89,” *Applied Economics*, vol. 28, no. 7, pp. 857–864, 1996.

[97] J. C. Smart, "Gender equity in academic rank and salary." *Review of Higher Education*, vol. 14, no. 4, pp. 511–26, 1991.

[98] B. D. Fisher, S. Motowidlo, and S. Werner, "Effects of gender and other factors on rank of law professors in colleges of business: Evidence of a glass ceiling," *Journal of Business Ethics*, vol. 12, no. 10, pp. 771–778, 1993.

[99] D. University of Californa, "Joint administration - academic senate task force on analysis of faculty salary equity at the university of california, davis," University of Californa , Davis, Tech. Rep., 2013. [Online]. Available: http://academicsenate.ucdavis.edu/local_resources/docs/Faculty%20Salary%20Equity/faculty-salary-equity-original-request.pdf

[100] E. L. Scott, "Higher education salary evaluation kit." 1977.

[101] J. C. Alessio, "Using discriminant analysis to predict professorial rank by gender," *Journal of Applied Sociology*, vol. 16, no. 1, pp. 123–143, 1999. [Online]. Available: http://www.jstor.org/stable/43481421

[102] C. F. G. Constantine Loucopoulos, Robert Pavur, "Insights into gender discrimination in employment compensation through the use of classification models," *Journal of Managerial Issues*, vol. 14, no. 3, pp. 375–387, 2002. [Online]. Available: http://www.jstor.org/stable/40604397

[103] I. Walters and N. le Roux, "Monitoring gender remuneration inequalities in

academia using biplots," *ORiON*, vol. 24, no. 1, 2008. [Online]. Available: http://orion.journals.ac.za/pub/article/view/59

[104] N. Moore, "Faculty salary equity: Issues in regression model selection," *Research in Higher Education*, vol. 34, no. 1, pp. 107–126, 1993.

[105] J. S. Osborne, "A strategy toward accountability: identifying faculty salary policies at kentucky public regional institutions," Ph.D. dissertation, University of Kentucky, 1990.

[106] N. Boudreau, J. Sullivan, W. Balzer, A. M. Ryan, R. Yonker, T. Thorsteinson, and P. Hutchinson, "Should faculty rank be included as a predictor variable in studies of gender equity in university faculty salaries?" *Research in Higher Education*, vol. 38, no. 3, pp. 297–312, 1997.

[107] M. W. Gray and E. L. Scott, "A" statistical" remedy for statistically identified discrimination," *Academe*, vol. 66, no. 4, pp. 174–181, 1980.

[108] R. McNabb and V. Wass, "Male-female salary differentials in british universities," *Oxford Economic Papers*, vol. 49, no. 3, pp. 328–343, 1997.

[109] D. A. Barbezat and J. W. Hughes, "Sex discrimination in labor markets: The role of statistical evidence: Comment," *The American Economic Review*, vol. 80, no. 1, pp. 277–286, 1990.

[110] W. Balzer, N. Boudreau, P. Hutchinson, A. M. Ryan, T. Thorsteinson, J. Sullivan, R. Yonker, and D. Snavely, "Critical modeling principles when testing for gender equity in faculty salary," *Research in Higher Education*, vol. 37, no. 6, pp. 633–658, 1996.

[111] L. A. Braskamp, J. A. Muffo, and I. W. Langston, "Determining salary equity: policies, procedures, and problems," *The Journal of Higher Education*, pp. 231–246, 1978.

[112] J. Pfeffer and J. Ross, "Gender-based wage differences the effects of organizational context," *Work and Occupations*, vol. 17, no. 1, pp. 55–78, 1990.

[113] R. T. Blackburn and J. H. Lawrence, *Faculty at Work. Motivation, Expectation, Satisfaction.* ERIC, 1995.

[114] D. Loomis, D. B. Richardson, and L. Elliott, "Poisson regression analysis of ungrouped data," *Occupational and Environmental Medicine*, vol. 62, no. 5, pp. 325–329, 2005. [Online]. Available: http://oem.bmj.com/content/62/5/325.abstract

[115] F. Y. Edgeworth, "Equal pay to men and women for equal work," *The Economic Journal*, vol. 32, no. 128, pp. 431–457, 1922.

[116] J. V. Koch and J. F. Chizmar, "Sex discrimination and affirmative action in faculty salaries," *Economic Inquiry*, vol. 14, no. 1, pp. 16–24, 1976.

[117] K. R. Browne, "Statistical proof of discrimination: beyond damned lies," *Wash. L. Rev.*, vol. 68, p. 477, 1993.

[118] G. J. Matthews and O. Harel, "Data confidentiality: A review of methods for statistical disclosure limitation and methods for assessing privacy," *Statistics Surveys*, vol. 5, pp. 1–29, 2011.

[119] G. W. Imbens and D. B. Rubin, "Rubin causal model," in *The New Palgrave Dictionary of Economics*, S. N. Durlauf and L. E. Blume, Eds. Basingstoke: Palgrave Macmillan, 2008.

[120] D. B. Rubin, "Estimating causal effects of treatments in randomized and non-randomized studies." *Journal of educational Psychology*, vol. 66, no. 5, p. 688, 1974.

[121] P. Diggle, S. Zeger, K. Liang, and P. Heagerty, *Analysis of longitudinal data*. Oxford University Press, 2002.

[122] N. Duan, "Smearing estimate: A nonparametric retransformation method," *Journal of the American Statistical Association*, vol. 78, no. 383, pp. 605–610, 1983. [Online]. Available: http://www.jstor.org/stable/2288126

[123] L. H. Moulton and S. L. Zeger, "Bootstrapping generalized linear models," *Computational Statistics & Data Analysis*, vol. 11, no. 1, pp. 53–63, 1991.

[124] M. A. Hussey and J. P. Hughes, "Design and analysis of stepped wedge cluster

randomized trials," *Contemporary Clinical Trials*, vol. 28, no. 2, pp. 182 – 191, 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1551714406000632

[125] K. Hemming, T. P. Haines, P. J. Chilton, A. J. Girling, and R. J. Lilford, "The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting," *BMJ*, vol. 350, 2015. [Online]. Available: http://www.bmj.com/content/350/bmj.h391

[126] AAMC, "2014 aamc annual report," Association of American Medical Colleges, Tech. Rep., 2012. [Online]. Available: https://members.aamc.org/eweb/upload/Diversity%20in%20Medical%20Education%20Facts%20and%20Figures%202012.pdf

[127] H. Alexander and C. Q. Liu, "The aging of full-time u.s. medical school faculty: 1967-2007," *Analysis in Brief*, vol. 9, no. 4, October 2009. [Online]. Available: https://www.aamc.org/download/102368/data/aibvol9no4.pdf

[128] T. H. J. K. F. Foundation, "Distribution of medical school graduates by gender," [Accessed 29-September-2015]. [Online]. Available: http://kff.org/other/state-indicator/medical-school-graduates-by-gender/

[129] AAMC, "Number of active residents by type of medical school graduation, gme specialty, and gender," [Accessed 29-September-2015]. [Online]. Available: https://www.aamc.org/data/421322/tableb2.html

[130] ——, "Total u.s. medical school graduates by race/ethnicity and sex, 2012-2013 and 2013-2014," [Accessed 04-October-2015]. [Online]. Available: https://www.aamc.org/download/321536/data/factstable29.pdf

[131] ——, "Diversity in medical education: Facts and figures 2012," Association of American Medical Colleges, Tech. Rep., 2015. [Online]. Available: https://members.aamc.org/eweb/upload/2014%20Annual%20Report%20non-flash.pdf

[132] J. Grazzini, "Analysis of the emergent properties: Stationarity and ergodicity," *Journal of Artificial Societies and Social Simulation*, vol. 15, no. 2, p. 7, 2012. [Online]. Available: http://jasss.soc.surrey.ac.uk/15/2/7.html

FRANCIS ABREU

615 N Wolfe Street, suite E3036
Baltimore, MD 21205
fabreu1@jhu.edu
http://francisabreu.github.io/

EDUCATION

2010-2016 **PhD in Biostatistics**. Johns Hopkins Bloomberg School of Public Health.
Advisor: Scott Zeger, PhD. Area of study: Analysis of clustered longitudinal data.

2004-2006 **Specialist in Statistics**. Universidad Central de Venezuela.
Advisor: Maura Vásquez, PhD. Area of study: Applied analysis of high dimensional data.

1998-2003 **Bachelor in Statistical Sciences**. Universidad Central de Venezuela.
Advisor: Rubén Ibarra. Area of Study: Applied econometric analysis.

RESEARCH

2014-present **Collaborator.** Johns Hopkins University School of Medicine, Dept. of Plastic and Reconstructive Surgery.

- Designed statistical plan for research project "Post-Operative Quality of Life Outcomes in Elderly Patients Who Underwent Mastectomy With or Without Breast Reconstruction" (submitted for IRB).
- Analyzed surgical outcomes using data from the National Surgical Quality Improvement Program (NSQIP) for several projects.

2011-present **Graduate research assistant**. Johns Hopkins Bloomberg School of Public Health, Dept. of Biostatistics.
Dissertation advisors: Scott Zeger, PhD and Elizabeth Colantuoni, PhD.

- Performed applied analyses and developed statistical methods to assess gender disparities in faculty trajectories.
- Developed applications to understand and visualize sources of information in stepped wedge designs.

2011-present **Graduate research assistant**. Johns Hopkins Bloomberg School of Public Health, Center for Clinical Trials.
Supervisors: Tom Louis, PhD and Elizabeth Sugar, PhD.

- Performed validation, maintained data grid freezes and documentation for the Multicenter Uveitis Steroid Treatment (MUST) Trial.
- Performed statistical analyses on ophthalmologic outcomes from the MUST Trial, in particular cataract surgery outcomes in uveitis and prevalence and risk factors of epiretinal membrane.

INDUSTRY / CONSULTING

2006-2010 **Senior Statistician and Project Manager.** Instituto Delphos, CA., Venezuela (ICRO).

- Corporate liaison between Delphos, pharmaceutical companies and medical researchers.
- Participated in the design of clinical studies, reviewed study protocols and CRFs with medical team, performed sample size calculations, and developed statistical analysis plans.
- Performed statistical programming and analysis, and delivered study reports.
- Participated and led presentation of results and workshops for medical researchers.

Fall 2008 **Statistical consultant**. Intelimedia, C.A., Venezuela.
Analyzed voter polarization in the 2008 Libertador District regional elections.

Summer 2006 **Statistical consultant**. CONAFIN Consultores, C.A., Venezuela.
Designed and evaluated questionnaires for client satisfaction surveys.

2003-2006 **Statistical Consultant**. United Nations Development Programme (UNDP).

- Developed statistical products and training activities to further local reality understanding, promotion of proper use of statistical information and strengthening of the National Health Statistical System.
- Conducted a diagnostic study of the Statistical System of the Ministry of Health (MSDS).
- Designed, organized and lead working groups of MSDS departments by priority health areas in order to analyze the relevance and flow of health information.

INDUSTRY / CONSULTING (CONT.)

Summer 2004 **Expert witness**. Arbitral Court TEL-FREE VENEZUELA vs. TELECOMUNICACIONES MOVILNET
Statistical considerations on goal achievements in the Sponsored Communications Program.

2002 **Statistical Intern**. Banco del Caribe, Dept. of Planning and Development, Venezuela.
Performed econometric analyses of office profitability determinants.

2000 **Statistical Intern**. Colegio Santiago de Leon de Caracas, Venezuela.
Performed statistical analyses of the Thinking Processes program Evaluation Survey.

TEACHING AND MENTORING

2011-2016 **Lead teaching assistant / Lab instructor.** Johns Hopkins Bloomberg School of Public Health.
Designed and lead labs with interactive exercises on theory, design, interpretation, computation and presentation of results. Held office hours and provided one-one consulting for final projects. Beta-tested, proctored and graded quizzes and exams.

- Analysis of Longitudinal Data (School-wide, ~ 90 enrollment). Spring 2013, 2014 & 2016.
- Statistical Methods in Public Health I-III (MPH-level, ~500 enrollment). Fall 2013, fall 2014 & spring 2015.
- Multilevel Statistical Models in Public Health (School-wide, ~70 enrollment). Spring 2013 & 2014.
- Essentials of Probability and Statistical Science I-IV (ScM-level, ~25 enrollment), 2011-2012 year long.

2014-2015 **Guest lecturer.** Johns Hopkins Bloomberg School of Public Health.

- Statistical Methods in Public Health II (MPH-level, ~500 enrollments). Spring 2015.
- Analysis of Longitudinal Data (School-wide, ~90 enrollment). Spring 2014.

2013-2014 **Visiting lecturer.** American University of Armenia.
Inferential Biostatistics (MPH-level, ~20 enrollment). Designed and delivered lectures and labs with interactive activities (theory, application and computing). Fall 2013 2014.

Spring 2014 **Lead teaching assistant.** Johns Hopkins Krieger School of Arts and Sciences
Health Data Analysis Practicum (Undergraduate, ~15 enrollment).

Summer 2013 **Visiting lecturer.** Universidad Central de Venezuela, Postgrad. Area in Statistics and Actuarial Sciences.
Survival Analysis (ScM-level, ~50 enrollment).

2012-2014 **Teaching assistant.** Johns Hopkins Bloomberg School of Public Health.

- Case-based Introduction to Biostats. (Coursera MOOC, ~25K enrollment). Spring 2014 & Summer 2013.
- Design of Clinical Experiments (School-wide, ~30 enrollment). Fall 2012.
- Statistical Reasoning in Public Health I (Online, ~200 enrollment). Fall 2012.

2011-2014 **Teaching assistant.** Johns Hopkins Bloomberg School of Public Health, Summer Institute of Epidemiology and Statistics.

- Longitudinal Data Analysis (~50 enrollment). Summer 2014.
- Data Analysis Workshop I & II (~60 enrollment). Summer 2014.
- Statistical Reasoning in Public Health I & II (~100 enrollment). Summer 2012 & 2013.
- Biostatistics in Medical Product Regulation (~10 enrollment). Summer 2011.

2008-2010 **Mentor**. Universidad Central de Venezuela, School of Statistics and Actuarial Sciences.
Mentored 7 students conducting applied research and developing statistical products for industry internships (commercial banks, pharmaceutical and marketing companies)

2007-2010 **Instructor (tenured 2010).** Universidad Central de Venezuela, School of Statistics and Actuarial Sciences.
Developed and taught ~ 4 undergraduate-level classes per semester with a co-instructor. Participated in oral examination committees (8 undergraduate, 1 ScM). Courses taught:

- Statistical Inference (5 semesters, ~50 enrollment).
- Sampling I (7 semesters, ~ 50 enrollment).
- Sampling II (3 semesters, ~ 50 enrollment).
- Data Analysis (4 semesters, ~20 enrollment).

Fall 2006 **Lead teaching assistant.** Universidad Central de Venezuela, School of Statistics and Actuarial Sciences.
Statistical Methods II (Undergraduate-level, ~40 enrollment).

Summer 2006 **Visiting lecturer.** Universidad Central de Venezuela, Postgrad. Area in Statistics and Actuarial Sciences.
Probability Theory and Sampling in Audit processes (Professional-level, ~40 enrollment).

Publications

1 Sen HN, **Abreu FM**, Louis TA, Sugar EA, Altaweel MM, Elner SG, Holbrook JT, Jabs DA, Kim RY, Kempen JH, Multicenter Uveitis Steroid Treatment (MUST) Trial and Follow-up Study Research Group. Cataract Surgery Outcomes in Uveitis: The Multicenter Uveitis Steroid Treatment Trial. *Ophthalmology*. 2016 Jan;123(1):183-90. doi: 10.1016/j.ophtha.2015.09.022. Epub 2015 Oct 20.

2 Abt NB, Flores JM, Baltodano PA, Sarhane KA, **Abreu FM**, Cooney CM, Manahan M, Stearns V, Makary M, Rosson GD. (2014). Neoadjuvant Chemotherapy and Short-term Morbidity in Patients Undergoing Mastectomy With and Without Breast Reconstruction. *JAMA surgery*, *149*(10), 1068-1076.

3 Sarhane KA, Flores JM, Cooney CM, **Abreu FM**, Lacayo M, Baltodano PA, Ibrahim Z, Alrakan M, Brandacher G, Rosson GD. (2013). Preoperative Anemia and Postoperative Outcomes in Immediate Breast Reconstructive Surgery: A Critical Analysis of 10,958 Patients from the ACS-NSQIP database. Plastic and Reconstructive Surgery–Global Open, 1(5), e30.

Submitted / Submitting soon

1 Bilal U, Cooper R, **Abreu FM**, Nau C, Francis M, Glass T. (2016) Do Social Protection Policies Mitigate the Association between Economic Growth and Mortality?.(*submitted*)

2 Lim LL, **Abreu FM**, Sugar EA, Burke AE, Altaweel MM, Rao PK, Holbrook JT, Elner SG, Stawell RJ, Kempen JH. (2016). Prevalence and Risk Factors of Epiretinal membrane in the Multicenter Uveitis Steroid Treatment Trial. (*submitting soon*)

3 Flores JM, **Abreu FM**, Chisolm M, Terplan M. (2016) Preoperative cigarette/alcohol use increases 30-day post-surgical mortality/morbidity and accelerates time to death. (*submitting soon*)

Published abstracts

1 Qadi MA, Baltodano PA, Flores JM, Reddy S, Abt NB, Sarhane KA, **Abreu FM,** Azih LC, Cooney CM, Rosson G. D. (2014). Are Flaps Really Better Than Implants for Breast Reconstruction in Obese Females? An Analysis of 89,514 Women Undergoing Breast Surgery from the ACS-NSQIP Database. *Plastic and reconstructive surgery*, *133*(4S), 982-983.

2 Baltodano PA, Flores JM, Kone L, Abt NB, Sarhane KA, Rochlin DH, **Abreu FM**, Zellars RC, Makary MA, Rosson GD. (2014). Neoadjuvant Radiotherapy is not associated with Increased Post-Mastectomy/Reconstruction Morbidity Events: A Critical Analysis of 85,851 Patients from the ACS-NSQIP Database. *Plastic and Reconstructive Surgery*, *133*(3s), 135-136.

3 Baltodano PA, Flores JM, Abt NB, Sarhane KA, **Abreu FM**, Burce KKl Cooney CMl Cooney DS, Sacks JM, Rosson G. D. (2014). Timing and Technical Implications of Breast Reconstruction in Anemic Women: The Advantages of Staged (Delayed-Immediate) Breast Reconstruction. *Plastic and Reconstructive Surgery*, *133*(3s), 111.

4 Abt NB, Flores JM, Baltodano P, Sarhane KA, Kone L, **Abreu FM**, Cooney CM, Makary MA, Rosson G. D. (2014). Abstract P24: Neoadjuvant Chemotherapy is Associated with Decreased Morbidity amongst 77,958 Patients Undergoing Mastectomy-only and Immediate Tissue Expander Reconstruction. *Plastic and Reconstructive Surgery*, *133*(3s), 205-206.

5 Sarhane KA, Flores JM, Shore AD, Baltodano PA, **Abreu FM**, Rosson G. D, Makary MA, Lee WP, Brandacher G, Sacks JM. (2014). Development and Validation of a Stratification Tool for Identifying Breast Cancer Patients with Elevated BMI at Increased Risk for Surgical Site Infections. *Journal of the American College of Surgeons*, *219*(3), S87.

6 Sarhane, KA, Flores JM, Shore AD, **Abreu FM**, Ibrahim Z, Alrakan M, Cooney CM, Baltodano PA, Drogt C, Makary MA, Brandacher G, Rosson G. D. (2013). A Validated, Risk Assessment Model for Predicting Morbidity after Breast Surgery. *Plastic and Reconstructive Surgery*, *132*(4S-1), 125.

7 Meléndez MM, Baltodano PA, Flores JM, Sarhane KA, **Abreu FM**, Rosson GD. (2013). Perioperative Transfusions and Postoperative Outcomes in Free Flap Reconstructive Surgery: A Critical Analysis of 6,132 Patients from the ACS-NSQIP Database. *Plastic and Reconstructive Surgery*, *132*(4S-1), 39.

8 Sarhane KA, Flores JM, **Abreu FM**, Ibrahim Z, Alrakan M, Cooney CM, Baltodano PA, Drogt C, Brandacher G, Rosson G. D. (2013). Building a clinical risk model to predict morbidity after breast surgery. *Journal of the American College of Surgeons*, *217*(3), S89.

PRESENTATIONS

2015 Flores JM, **Abreu FM**, Chisolm M, Terplan M. Preoperative cigarette/alcohol use increases 30-day post-surgical mortality/morbidity and accelerates time to death. College on Problems of Drug Dependence - 77th Annual Meeting, Phoenix, USA. (Podium)

2015 Lim LL, **Abreu FM**, Sugar EA, Burke A, Altaweel MM , Rao PK , Holbrook JT, Elner SG , Stawell R, Kempen, JH.Prevalence and Risk Factors of Epiretinal membrane in the Multicenter Uveitis Steroid Treatment (MUST) Trial. The Association for Research in Vision and Ophthalmology Annual Meeting 2015, Denver, USA. (Poster)

2014 Sen HN, **Abreu FM**, Louis TA, Sugar E, Altaweel MM, Elner SG, Holbrook JT, Jabs DA, Kim RY, Kempen J. Cataract Surgery Outcomes in Uveitis: The Multicenter Uveitis Steroid Treatment (MUST) Trial. American Academy of Ophthalmology Annual Meeting 2014, Chicago, USA. (Poster)

2010 **Abreu FM**, Torrealba A. Analysis of Air Pollution Statistics and Human Development, by Means of Principal Components and Cluster Analysis. IAOS 2010, Santiago de Chile, Chile. (Podium)

2007 **Abreu FM**, Vásquez ME. Analysis of interrelations between human development and socio-demographic environment by means of Canonical Correspondence Analysis. IV Meeting of the Central American and Caribbean Region of the International Biometric Society, Isla de Margarita – Venezuela. (Podium)

2005 Ramírez G, Vásquez ME, **Abreu FM**. An application of canonical correspondence analysis to diagnose the socioeconomic development of Miranda State. LV Annual Meeting of ASOVAC (Venezuelan Association for the Advancement of Science), Caracas – Venezuela. (Poster)

2004 **Abreu FM**, Proyecto SistEEM team. Gathering of information at Municipal Micro Areas level: pilot experience with the Public Health sector. II Meeting on Experiences on Local and Community Measurements for Micro-level Planning and Design of Public Policies, Ministry of Science and Technology, Caracas – Venezuela. (Podium)

2004 **Abreu FM**. Graduates in Statistics describe their job experiences. VI Academic Sessions of the Department of Statistics and Actuarial Sciences, Universidad Central de Venezuela, Caracas – Venezuela.

2002 **Abreu FM**. Talking to Academic Merit Award winning students. Informative sessions on Admissions to FACES, School of Economy and Social Sciences (FACES), Universidad Central de Venezuela, Caracas – Venezuela. (Podium)

CERTIFICATES – RESEARCH AND COMPUTING

2015 *Training/Practice of Unlicensed Clinical Research Staff*. Johns Hopkins Bloomberg School of Public Health.

2013 *Conflict of interest and Commitment*. Johns Hopkins Bloomberg School of Public Health.

HIPAA & Research. Johns Hopkins Bloomberg School of Public Health.

Intermediate Privacy course for Health Care Providers. Johns Hopkins Bloomberg School of Public Health.

Human Subjects Research. Johns Hopkins Bloomberg School of Public Health.

Business Ethics Training for Faculty and Staff Working on Federal Contracts. Johns Hopkins Bloomberg School of Public Health.

2009 *Corporate Activity and Human Rights*. Amnesty International – StatoilHydro, Venezuela.

2007 *Good Clinical Practices / Serious Adverse Events*. Novartis Pharma, Venezuela.

2002 *Geographical Information Systems*. Universidad Central de Venezuela.

CERTIFICATES – TEACHING AND LEADERSHIP

2014-2015 *Preparing Future Faculty Teaching Academy*. Johns Hopkins Bloomberg School of Public Health.

Participated in monthly seminars and workshops and completed the following activities:

- Phase I: An introduction to Evidence-Based Undergraduate STEM Teaching, Coursera MOOC.
- Phase II: Teaching at the University Level course.
- Phase III: Redesigned and taught Inferential Biostatistics course for the American University of Armenia.

2006 *Virtual Tutor Training*. Center for Virtual Distance Studies (CEVAD).

2004 *Teamwork and Leadership Techniques*. Latin American Women's Rights Foundation (FUNDEMUL), Venezuela.

ADDITIONAL PROFESSIONAL DEVELOPMENT

2015 *Research Leadership*. Johns Hopkins Bloomberg School of Public Health.

2006 *Statistical Consulting*. University of Wisconsin-La Crosse.

HONORS AND AWARDS - RESEARCH

2006 **Distinction of excellence** on specialization thesis. Universidad Central de Venezuela.

HONORS AND AWARDS - TEACHING

2015 **Helen abbey Award.** Johns Hopkins Bloomberg School of Public Health, Department of Biostatistics. Honors a Biostatistics student who has shown a commitment to teaching after graduation.

2011 **Eponym of the LXXVIII class of Statisticians and Actuaries**. Universidad Central de Venezuela. Honors teachers who students feel have contributed in a special way in their education.

2010 **Godmother to the LXXVI class of Statisticians and Actuaries**. Universidad Central de Venezuela. Honors teachers who students feel have contributed in a special way in their education.

HONORS AND AWARDS – ACADEMIC MERIT

2003 **Summa Cum Laude.** Bachelor in Statistical Sciences. Universidad Central de Venezuela. First student to receive this honor in the 50 years since the founding of the School of Statistics and Actuarial Sciences.

2003 **Best GPA in graduating class** of the School of Economic and Social Sciences II-2002. Universidad Central de Venezuela.

1999-2003 **Academic Merit Scholarship**. Universidad Central de Venezuela.
Best GPA in the Department of Statistics and Actuarial Sciences 1999-2003.
7th best overall GPA in the university 2002-2003.
5th best overall GPA in the university 2000-2001.

2002 **José Félix Ribas Order: Third degree, Academic Merit**. Ministry of Health and Social Development, Venezuela.

1997-2001 **Ernesto Rivas González Academic Merit Award**. Universidad Central de Venezuela. Best GPA in the Department of Statistics and Actuarial Sciences for all eligible semesters (9)

1999 **Study Tour Awards for Outstanding Students of the Japanese Language**. The Japan Foundation Japanese Language Institute, Japan.

SERVICE

2015-present **Reviewer** for Annals of Surgical Oncology

Spring 2014 **Volunteer** ENAR Spring meeting. Baltimore, MD.

2013-2014 **Student Grievance Board Member**. Johns Hopkins Bloomberg School of Public Health.

2012-2014 **TA training day coordinating committee member**. Johns Hopkins Bloomberg School of Public Health. Dept. of Biostatistics.

2005-2010 **Secretary of External Affairs**. Venezuelan Society of Statisticians and Actuaries.

MEMBERSHIP IN PROFESSIONAL SOCIETIES

2014-present Institute of Mathematical Statistics.

2004-present Venezuelan Society of Statisticians and Actuaries.

2004-present Venezuelan College of Statisticians and Actuaries.

Languages

Spanish (native), English (bilingual), Japanese (basic), French (basic)

Software skills

MS Office: Word, Excel, Access, Power Point, Project, OneNote, among others
Statistical: R, STATA, SAS, SPSS, Minitab, Eviews, SPAD, Redatam, among others
Markup: TeX, LaTeX, Beamer, TeXworks
Programming: Python, C++ (basic)

References

Scott Zeger, PhD. sz@jhu.edu.
Professor of Biostatistics (thesis advisor). Johns Hopkins Bloomberg School of Public Health.
Elizabeth Colantuoni, PhD. ejohnso2@jhmi.edu.
Associate Scientist in Biostatistics (thesis co-advisor). Johns Hopkins Bloomberg School of Public Health.
Tom Louis, PhD. tlouis1@jhu.edu.
Professor of Biostatistics (academic advisor and collaborator). Johns Hopkins Bloomberg School of Public Health.